



---

# HETEROCEDASTICIDAD CONDICIONAL, ATÍPICOS Y CAMBIOS DE NIVEL EN SERIES TEMPORALES FINANCIERAS

---

TESIS DOCTORAL

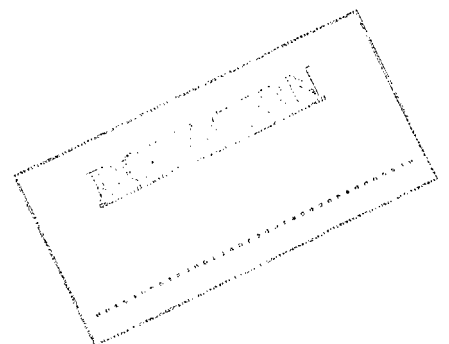
**Autora:** María Angeles Carnero Fernández

**Directores:** Daniel Peña y Esther Ruiz



UNIVERSIDAD CARLOS III DE MADRID  
Departamento de Estadística y Econometría

Getafe, Enero de 2003



*A mis padres*

*Estudia el pasado si quieres pronosticar el futuro*

Confucio

## Agradecimientos

En esta Tesis han participado muchas personas, sin las cuales nunca habría sido posible. A todas ellas les debo mi gratitud.

Agradezco muy sinceramente a mis directores, Daniel Peña y Esther Ruiz, por su confianza y apoyo. Ellos son quien más de cerca han compartido conmigo este proyecto, ayudándome en todo momento, tanto a nivel profesional como personal. Sin duda, es gracias a ellos que he llegado hasta aquí.

Quisiera mostrar mi gratitud a todos los miembros del Departamento de Estadística y Econometría y en particular, a Francisco Javier Prieto, a Regina Kaiser, a Jesús Gonzalo y a Olivier Núñez. A Rosa Lillo le agradezco su amistad, ser su ayudante ha sido un placer. Gracias a Ismael Sánchez y a Felipe M. Aparicio por todos sus comentarios tras una lectura detallada de este trabajo.

Agradezco al Ministerio de Ciencia y Tecnología y a la sociedad, en general, por la financiación que me ha permitido llevar a cabo esta Tesis.

Quisiera mostrar mi agradecimiento a los profesores Carlos Hervés, por ser quien me embarcó en esta aventura e Ignacio García Jurado por su ayuda. A los profesores Granville Tunnicliffe-Wilson, Ruey Tsay, Richard Davis, Marco Reale, Les Oxley, Michael McAleer, Ron Bewley y Neil Shephard les agradezco su afectuosa acogida durante mis visitas respectivas a las Universidades de Lancaster, Chicago, Colorado, Canterbury, Western Australia, New South Wales y Oxford. De todos ellos he aprendido muchísimo.

A Luis Tenorio le agradezco el tiempo que compartimos, así como sus sabios consejos.

Mis compañeros de Doctorado han sido, sin duda, una parte importantísima de este proyecto, sin su amistad no habría llegado hasta aquí. Gracias a Andrés Alonso por la ayuda que me ha brindado tantas veces, a Andrés Ubierna por su alegría, a Oscar Martínez por sus múltiples enseñanzas y su portátil, tan útil ahora al final, a Alfonso Sánchez por sus conversaciones, a Carlos Bethencourt por su sentido del humor, a José María Carretero por tantos buenos ratos, a Pilar Poncela por ser un

ejemplo a seguir. Gracias también a Ana García, a Carmen Broto, a Pedro Galeano, a Ramiro Losada y a Ana Pérez. Gracias a Lorenzo Pascual por todo el ánimo y alegría que me ha dado durante estos años.

Con Rosa Loveira, Virginia Sánchez y Nieves Carrera he compartido muchos momentos buenos. Gracias por escucharme tantas veces y mostrarme, con vuestro ejemplo, lo mejor de las personas. Les agradezco todo el tiempo que me han dedicado. Han sido las mejores compañeras y amigas.

No tengo palabras de agradecimiento para Rocío Sánchez, más compañera del alma que de despacho. Con ella he compartido los mejores momentos de esta Tesis y su *¿sabes qué te digo?* me ha hecho reír en los peores. Me llevo su amistad como uno de los tesoros más preciados.

A Domingo Ortega debería ponerle en los desagradecimientos, por todo el tiempo que, gracias a él, le he quitado a esta Tesis y dedicado al teatro. Gracias por haber sido el mejor profesor. Gracias también a Sunny por su encantadora sonrisa. Gracias a Isabel Campo, mi amiga del alma, por haber estado tan cerca, estando tan lejos. Gracias a Mónica Pereira por las aventuras que compartimos. Gracias a Teri, Jenni, Nadir, Suhejla, Mahendra, Clinton y Roos. Gracias a todos por vuestra maravillosa amistad.

A mi familia se lo debo todo. Sin ellos, nada de esto habría sido posible. A mis padres les doy las gracias por su cariño y apoyo incondicional desde el principio, así como a mis hermanos Eli, Eduardo y Tono y a mis cuñadas Susana y Marilyn. Sentirles tan cerca a todos es algo que me enorgullece enormemente. A mi sobrino, Xabier, le doy las gracias por su cariño tan sincero y la alegría que ha traído a nuestras vidas. Gracias también a Conchita y a Paula por su cariño y todo el tiempo que me han dedicado. No tengo palabras para agradecer a todos mis primos, en especial a José Antonio y Manola, el cariño y afecto que me han dado durante todos estos años, así como a mi tía Concha. Gracias a Conchita, Jose, Javi, Nani, Fredi, Tareixa y Patricia por todos los días que compartisteis conmigo. Los domingos serán siempre un día maravilloso gracias a todos vosotros.

# Índice General

<b>1</b>	<b>Introducción</b>	<b>1</b>
<b>2</b>	<b>Modelos de heterocedasticidad condicional</b>	<b>7</b>
2.1	Características empíricas de los rendimientos financieros . . . . .	8
2.2	Modelos ARCH . . . . .	11
2.2.1	Modelos GARCH . . . . .	12
2.2.2	Modelos EGARCH . . . . .	20
2.2.3	Otros modelos ARCH para varianzas condicionales asimétricas .	23
2.3	Modelos de Volatilidad Estocástica . . . . .	25
2.4	Otros modelos alternativos . . . . .	31
2.5	Apéndice de tablas y figuras . . . . .	33
<b>3</b>	<b>Persistencia y curtosis en modelos GARCH y ARSV</b>	<b>51</b>
3.1	Introducción . . . . .	52
3.2	Propiedades de los modelos GARCH y ARSV . . . . .	54
3.2.1	El modelo GARCH(1,1) . . . . .	54

---

3.2.2	El modelo EGARCH(1,1) . . . . .	56
3.2.3	El modelo ARSV(1) . . . . .	57
3.3	Experimentos de Monte Carlo . . . . .	60
3.4	Conclusiones . . . . .	62
3.5	Apéndice de tablas y figuras . . . . .	64
<b>4</b>	<b>Observaciones atípicas y heterocedasticidad condicional</b>	<b>73</b>
4.1	Introducción . . . . .	74
4.2	Observaciones atípicas en series temporales . . . . .	75
4.3	Efectos en la identificación de la heterocedasticidad condicional . . . . .	79
4.3.1	Contrastes de homocedasticidad condicional . . . . .	79
4.3.2	Sesgos en los correlogramas . . . . .	82
4.4	Efectos en la estimación de la heterocedasticidad condicional . . . . .	87
4.4.1	Estimación de modelos ARCH . . . . .	87
4.4.2	Sesgos en los estimadores MCO de los modelos ARCH( $p$ ) . . . . .	90
4.4.3	Sesgos en los estimadores MV . . . . .	99
4.5	Modelización de la heterocedasticidad condicional en presencia de ob- servaciones atípicas . . . . .	103
4.6	Ejemplos . . . . .	106
4.7	Conclusiones . . . . .	110
4.8	Apéndice de tablas y figuras . . . . .	111
<b>5</b>	<b>Detección de cambios de nivel</b>	<b>129</b>

5.1	Introducción . . . . .	130
5.2	Método de Tsay para detectar cambios de nivel . . . . .	132
5.3	Un contraste alternativo. Distribución asintótica y propiedades en muestras finitas . . . . .	136
5.4	Detección de un cambio de nivel en modelos ARIMA . . . . .	141
5.5	Detección de múltiples cambios de nivel . . . . .	142
5.6	Aplicación empírica . . . . .	146
5.7	Conclusiones . . . . .	148
5.8	Apéndice A . . . . .	149
5.9	Apéndice de tablas y figuras . . . . .	150
6	Consideraciones finales y líneas de investigación futura	163



# Capítulo 1

## Introducción

La econometría financiera constituye actualmente un área de investigación muy activa. Hasta hace veinte años, la mayoría de los artículos empíricos en finanzas utilizaban herramientas analíticas muy simples. Sin embargo, el rápido desarrollo de la tecnología informática, la disponibilidad de nuevos datos, junto con el desarrollo y adaptación de técnicas econométricas más sofisticadas han cambiado el área por completo. Además, estos avances no se han restringido a la investigación académica sino que han tenido también una gran influencia en la práctica de las finanzas. El desarrollo de la econometría financiera ha tenido, fundamentalmente, dos objetivos. Por una parte, proponer modelos que representen las propiedades empíricas más habituales de los datos financieros. Por otra, desarrollar procedimientos para la estimación y contraste de los modelos propuestos por la Teoría Financiera. Esta Tesis se centra en el primer objetivo y, en particular, en el análisis de algunas propiedades dinámicas de series temporales financieras univariantes observadas con frecuencias elevadas como, por ejemplo, semanal o diariamente. Se analizan tres características empíricas frecuentemente observadas en los rendimientos financieros: (1) heterocedasticidad condicional, que implica exceso de curtosis y dependencia temporal en los segundos momentos de la serie con correlaciones pequeñas aunque muy persistentes, (2) observaciones atípicas, que son

aquellas observaciones provocadas por fenómenos extraordinarios que no forman parte de la dinámica habitual de la serie, y (3) cambios de nivel, que son cambios en la media marginal de la serie. Las principales aportaciones de esta Tesis se encuentran en estas tres direcciones.

En el Capítulo 2 se hace una revisión actualizada de la literatura existente sobre los modelos de heterocedasticidad condicional más utilizados en la práctica para representar series de rendimientos financieros. Se analizan detalladamente los modelos GARCH (Generalized AutoRegressive Conditional Heterocedasticity), EGARCH (Exponential GARCH) y ARSV (AutoRegressive Stochastic Volatility). La diferencia entre estos modelos estriba, principalmente, en que los dos primeros modelizan la volatilidad como la varianza condicional de la serie temporal, es decir, como una función de observaciones pasadas mientras que los ARSV modelizan la volatilidad como un proceso estocástico con un componente predecible y otro inesperado. En este capítulo se muestra que, tanto los modelos GARCH como los ARSV son capaces de representar adecuadamente series con exceso de curtosis y cuyos cuadrados están correlacionados. Sin embargo, aunque hasta el momento se había creído que las funciones de autocorrelación de las observaciones al cuadrado de ambos modelos eran muy similares, se muestra que puede haber diferencias importantes en la estructura de dichas funciones. Además, se ilustra que, cuando ambos modelos se ajustan a la misma serie temporal, la persistencia de la volatilidad estimada con los modelos GARCH es mayor que con los modelos ARSV. También se muestra que los modelos ARSV gaussianos parecen ser adecuados a la hora de explicar las propiedades de las series reales, mientras que si el modelo que se ajusta es un GARCH, la distribución condicional necesita ser leptocúrtica.

En el Capítulo 3 se explica por qué la persistencia estimada en los modelos GARCH es mayor que la estimada en ARSV y por qué el supuesto de gaussianidad condicional es adecuado para los ARSV pero no lo es para los GARCH. Estos hechos ya habían

sido observados, pero no explicados, en la literatura. Para ello se analiza la relación entre el coeficiente de curtosis marginal, la persistencia de las innovaciones sobre la volatilidad y la correlación de orden uno de las observaciones al cuadrado de los modelos GARCH(1,1) y ARSV(1). Utilizando dicha relación, es posible observar que los modelos GARCH sólo son capaces de generar series con exceso de curtosis y correlaciones de los cuadrados de magnitudes similares a las observadas en la práctica, forzando la persistencia de la volatilidad hacia la frontera de la raíz unitaria. Otra alternativa es introducir una distribución condicional con colas pesadas, que pueda explicar la curtosis sin incrementar excesivamente las correlaciones de los cuadrados. Sin embargo, en este caso, es frecuente encontrar modelos donde la condición de existencia del momento de cuarto orden no se satisface. Alternativamente, la introducción de un error adicional en la volatilidad permite que los modelos ARSV puedan generar series con curtosis grandes y correlaciones de los cuadrados pequeñas para un rango más amplio de parámetros. En este sentido, se verá que los modelos ARSV son más flexibles que los modelos GARCH para representar las propiedades empíricas que se observan habitualmente en los rendimientos financieros.

En el Capítulo 4 se estudia el problema que plantea la presencia de observaciones atípicas en series temporales condicionalmente heterocedásticas. Como se ha comentado anteriormente, introducir distribuciones condicionales con colas anchas no permite explicar totalmente la elevada curtosis frecuentemente observada en los residuos de los modelos de heterocedasticidad condicional, la cual podría estar causada por la existencia de observaciones atípicas. La presencia simultánea en una serie temporal de heterocedasticidad condicional y observaciones atípicas es un problema complicado dado que los efectos de ambos fenómenos pueden confundirse. Ambos causan exceso de curtosis y, si los atípicos son consecutivos, correlaciones en los cuadrados. Además, las observaciones correspondientes a etapas de volatilidad alta podrían ser identificadas como atípicos si se utilizan criterios tradicionales basados en la varianza marginal. Por

otra parte, las observaciones atípicas siempre pueden interpretarse como observaciones generadas por una distribución de mayor varianza que la del resto de las observaciones.

En concreto, en este capítulo se estudian los efectos causados por la presencia de observaciones atípicas en la identificación de la heterocedasticidad condicional y se muestra que un atípico aislado sesga hacia cero todos los coeficientes de correlación de las observaciones al cuadrado, y rachas de atípicos generan correlaciones en los cuadrados de la serie, lo que puede llevar a una detección espuria de heterocedasticidad condicional. A continuación, se derivan los sesgos asintóticos causados por los atípicos en los estimadores de Mínimos Cuadrados Ordinarios (MCO) de procesos ARCH( $p$ ) encontrando que atípicos aislados llevan a infraestimar los parámetros que miden la heterocedasticidad y atípicos consecutivos pueden sobreestimarlos. Mediante estudios de simulación se ilustran tales efectos en muestras finitas, tanto en los estimadores MCO como en los estimadores de Máxima Verosimilitud (MV) de modelos ARCH( $p$ ) y GARCH(1,1). Finalmente, en este capítulo, se propone un procedimiento, en tres pasos, para tratar la presencia simultánea de observaciones atípicas y heterocedasticidad condicional en series temporales. Dicho procedimiento consiste en: (i) detectar y corregir los atípicos más extremos con objeto de evitar los sesgos mayores en una primera estimación de las varianzas condicionales, (ii) calcular las observaciones estandarizadas, detectar y corregir los atípicos restantes en dichas observaciones siguiendo los métodos tradicionales para series homocedásticas y, (iii) reestimar las varianzas condicionales.

El efecto de las observaciones atípicas consideradas en el Capítulo 4, sobre la dinámica de la serie es transitorio. Sin embargo, al estudiar series por periodos largos de tiempo, es posible encontrar acontecimientos extraordinarios con efectos permanentes que provoquen, por ejemplo, cambios de nivel. El Capítulo 5 aborda el problema de la detección de dichos cambios. En primer lugar, se analizan las propiedades de un estadístico comúnmente usado para detectar cambios de nivel en series temporales. Se muestra que la distribución asintótica de dicho estadístico no es conocida y su dis-

tribución en muestras finitas depende crucialmente del tamaño muestral y del proceso generador de los datos. A continuación, se propone un nuevo contraste para el que se deriva su distribución asintótica. Mediante experimentos de Monte Carlo se analizan sus propiedades en muestras finitas y se muestra que la distribución asintótica proporciona una buena aproximación de dichas propiedades incluso para tamaños de muestra moderados. Además, la distribución muestral del nuevo estadístico no se ve afectada por el hecho de que la serie para la que se hace el contraste esté generada por un proceso con heterocedasticidad condicional. Por lo tanto, la distribución asintótica obtenida es una buena aproximación a la distribución muestral, al menos, en el contexto de las series temporales condicionalmente heterocedásticas de interés en esta Tesis. Finalmente, en este capítulo, se analiza también el problema de la detección de múltiples cambios de nivel en una serie. Se propone un procedimiento que consiste en detectar un cambio de nivel, dividir la serie en las dos subseries correspondientes y, buscar nuevos cambios en cada una de las subseries. Se muestra que este procedimiento aventaja a otro, más utilizado en la práctica, que consiste en detectar un cambio de nivel, corregir la serie por su magnitud y buscar cambios adicionales en la serie corregida. La ventaja fundamental del procedimiento propuesto es que su objetivo es detectar todos los cambios existentes en la serie antes de estimar sus respectivos tamaños lo cual evita sesgos en la estimación y, consecuentemente, la generación de cambios de nivel espurios, que son los principales problemas del procedimiento tradicional.

Para ilustrar todos los resultados obtenidos, a lo largo de la Tesis se analizan diferentes series diarias de rendimientos financieros entre los que se encuentran varios tipos de cambio y algunos índices bursátiles.

Por último, el Capítulo 6 contiene las consideraciones finales y líneas de investigación futuras.

Las principales aportaciones de la Tesis se pueden resumir en los siguientes puntos:

1. Clarificar las diferencias entre modelos GARCH y ARSV, lo que permite explicar algunos fenómenos observados en análisis empíricos previos de series diarias de rendimientos financieros.
2. Obtener los sesgos causados por atípicos aditivos en el correlograma de los cuadrados y en los estimadores de modelos ARCH.
3. Proponer un procedimiento para modelizar la presencia simultánea de observaciones atípicas y heterocedasticidad condicional en series temporales.
4. Analizar las propiedades del estadístico de Tsay para detectar cambios de nivel bajo heterocedasticidad condicional y mostrar que dicho estadístico no es robusto a la presencia de heterocedasticidad condicional.
5. Proponer un nuevo contraste para detectar cambios de nivel, que mejora al anterior, y obtener su distribución asintótica.
6. Proponer una nueva metodología para tratar el problema de múltiples cambios de nivel.

## Capítulo 2

# Modelos de heterocedasticidad condicional

El objetivo de este capítulo es hacer una revisión de la literatura sobre los modelos de heterocedasticidad condicional más utilizados en la práctica para representar series de rendimientos financieros observadas con frecuencias elevadas. Se analizan, en detalle, los modelos GARCH, EGARCH y ARSV. Las características de estos modelos y los problemas que surgen habitualmente, cuando éstos se ajustan a series reales, se ilustran con doce series diarias de rendimientos financieros.

## 2.1 Características empíricas de los rendimientos financieros

La mayoría de los trabajos en finanzas usan series de rendimientos en lugar de las series originales de precios. Campbell et al. (1997) argumentan que hay al menos dos razones para ello. La primera es que, para el inversor medio, el rendimiento de un activo financiero es un indicador, libre de escala, de las oportunidades de inversión. La segunda, que las series de rendimientos tienen propiedades estadísticas más simples que las series de precios, como son la estacionariedad y la ergodicidad.

Considérese un activo financiero cuyo precio en el instante  $t$  es  $p_t$ . Habitualmente, el rendimiento de este activo se define como  $r_t = \frac{p_t - p_{t-1}}{p_{t-1}}$ . No obstante, es común aproximar este rendimiento por  $y_t = \log(p_t) - \log(p_{t-1})$  basándose en una expansión de Taylor de segundo orden de la función  $\log\left(\frac{p_t}{p_{t-1}}\right)$  alrededor de 1, ya que, por lo general,  $p_t$  es muy cercano a  $p_{t-1}$ , esto es,

$$\log\left(\frac{p_t}{p_{t-1}}\right) = \log(1) + 1\left(\frac{p_t}{p_{t-1}} - 1\right) + \frac{f(t)^2}{2}$$

donde  $f(t)^2$  es un término despreciable. Por tanto, la serie de interés viene dada por

$$y_t = \log(p_t) - \log(p_{t-1}) \approx \frac{p_t - p_{t-1}}{p_{t-1}} = r_t$$

Una razón alternativa para usar  $y_t$  como medida del rendimiento de un activo puede encontrarse en Törnqvist et al. (1985), quienes caracterizan la función logarítmica y sugieren que los cambios logarítmicos, es decir, la serie  $y_t$ , deberían ser los usados para medir cambios relativos o diferencias en las variables económicas porque el cambio logarítmico es el único indicador simétrico, aditivo y normado de los cambios relativos.

Las series de rendimientos financieros observadas con frecuencias elevadas están caracterizadas principalmente por tener distribuciones marginales leptocúrticas y correlaciones en los cuadrados. Estas propiedades fueron documentadas originalmente



por Mandelbrot (1963) y Fama (1965). Las series de rendimientos son habitualmente incorreladas pero no independientes. La función de autocorrelación (fac) de los rendimientos al cuadrado tiene un coeficiente de primer orden normalmente pequeño y positivo seguido por coeficientes que decaen hacia cero muy lentamente; ver, por ejemplo, Bollerslev y Engle (1993) y Mills (1996). Es importante notar que estas propiedades dependen de la frecuencia con se observa la serie y parecen ser diferentes si se consideran series mensuales; véase, por ejemplo, Aparicio y Estrada (2001).

Con objeto de ilustrar dichas propiedades, la tabla 2.1 contiene estadísticos descriptivos para las series diarias de los rendimientos, en porcentaje, del dólar estadounidense frente al dólar canadiense, la peseta española, el marco alemán, el yen japonés, el franco suizo, la corona sueca y la libra esterlina observados desde enero de 1993 hasta octubre de 2000<sup>1</sup>. Se describen también cinco índices bursátiles internacionales, los de Amsterdam y Bombay (observados diariamente desde octubre de 1995 a octubre de 2000), el Dow Jones (desde enero de 1990 a octubre de 2000), el IBEX 35 (desde enero de 1992 a diciembre de 1999) y, finalmente, el S&P 500 (desde noviembre de 1987 a diciembre de 1998). La tabla 2.1 muestra que la mayoría de estas series tienen media cero y todas ellas tienen exceso de curtosis y coeficientes de simetría negativos. También puede observarse que, aunque la mayoría de las series son incorreladas, no son independientes ya que los rendimientos al cuadrado tienen correlaciones significativas. Las figuras 2.1 y 2.2, que representan las series de rendimientos de los tipos de cambio y de los índices bursátiles respectivamente, muestran agrupamientos de la volatilidad con secuencias de días con rendimientos grandes en valor absoluto. Ambas figuras contienen además estimadores kernel de las correspondientes densidades marginales junto con la densidad normal. Estos gráficos confirman que las distribuciones de los rendimientos financieros son leptocúrticas. Finalmente, se muestra la fac de las observaciones al cuadrado  $y_t^2$ ,

---

<sup>1</sup>Los datos se han obtenido en la página web <http://pacific.commerce.ubc.ca/xr/> por cortesía del Prof. Werner Antweiler, University of British Columbia, Vancouver, Canada.

donde las primeras correlaciones son pequeñas pero significativas incluso para retardos grandes. Este hecho puede sugerir la presencia de persistencia alta o memoria larga en la volatilidad del proceso; ver, por ejemplo, Ding et al. (1993), Bollerslev y Mikkelsen (1996), Lobato y Savin (1998) y Lobato y Velasco (2000).

Finalmente, la figura 2.3 representa, como ejemplo, la fac de la serie de rendimientos del S&P500,  $y_t$ , junto con las fac de  $y_t^2$ ,  $|y_t|$  y  $\log(y_t^2)$ . Dicho gráfico ilustra lo que es conocido como *efecto Taylor*, es decir, los rendimientos, en valor absoluto, tienen las mayores correlaciones de todas las posibles transformaciones exponenciales. Este resultado también ha sido encontrado por Taylor (1986), Cao y Tsay (1992), Ding et al. (1993) y Granger y Ding (1995), entre otros.

El modelo más simple capaz de representar las propiedades empíricas descritas, especifica la serie de interés,  $y_t$ , como el producto de dos procesos,  $\varepsilon_t$  y  $\sigma_t$ , esto es

$$y_t = \varepsilon_t \sigma_t \quad t = 1, 2, \dots, T \quad (2.1)$$

donde  $\varepsilon_t$  es un proceso ruido blanco independiente e idénticamente distribuido (i.i.d.) de media cero y varianza uno, independiente del proceso  $\sigma_t$ , conocido como volatilidad en la literatura financiera. La volatilidad es un factor crucial en muchos modelos de valoración de opciones y de formación de carteras. Su modelización sirve para calcular, por ejemplo, el valor en riesgo, así como para mejorar la precisión de los intervalos de predicción para los rendimientos futuros, más estrechos en periodos de calma financiera que en aquellos periodos más volátiles.

A pesar de que la volatilidad no es directamente observable, tiene ciertas características que son comúnmente observadas en las series de rendimientos financieros: (i) agrupamiento: la volatilidad suele ser grande en ciertos periodos de tiempo y pequeña en otros periodos, (ii) continuidad: la volatilidad evoluciona en el tiempo de un modo continuo, los cambios bruscos no son frecuentes, (iii) acotación: la volatilidad no diverge hacia infinito sino que varía en un rango finito, y (iv) asimetría: la volatilidad

parece reaccionar de un modo diferente a un incremento grande de los precios o a una gran caída. Todas estas propiedades juegan un papel muy importante en los modelos de volatilidad.

Es fácil ver que el proceso  $y_t$ , en (2.1), es una diferencia de martingalas, es decir,  $y_t$  tiene media cero, varianza marginal constante y condicionado al pasado, su esperanza es cero. En particular,  $y_t$  es incorrelado. Además, el modelo (2.1) puede generar correlaciones en los cuadrados a través de una dinámica autorregresiva en  $\sigma_t^2$ . El exceso de curtosis de los rendimientos,  $y_t$ , puede obtenerse de cualquiera de las maneras siguientes:

1. Colas pesadas en la distribución marginal de  $\varepsilon_t$ .
2. Heterocedasticidad condicional. La varianza condicional de  $y_t$  viene dada por  $E(y_t^2 | Y_{t-1}) = E(\sigma_t^2 | Y_{t-1})$  donde  $Y_{t-1} = (y_1, \dots, y_{t-1})$ , y el exceso de curtosis de  $y_t$  depende de la evolución dinámica de su varianza condicional.
3. Cambios aleatorios en la varianza y valores atípicos.

En la literatura se ha propuesto una amplia variedad de modelos para representar la evolución dinámica de  $\sigma_t$ . A lo largo de este capítulo se describirán algunos de ellos, tratando de seleccionar los que han recibido más atención en la práctica.

## 2.2 Modelos ARCH

Los modelos ARCH (AutoRegressive Conditional Heteroscedasticity) fueron propuestos originalmente por Engle (1982) para modelizar la varianza condicional de la inflación del Reino Unido. La idea básica de estos modelos es que la volatilidad,  $\sigma_t$ , es una función no lineal de las observaciones pasadas. Por ejemplo, en un modelo ARCH(2), la volatilidad  $\sigma_t$  en (2.1) viene determinada por  $\sigma_t^2 = \omega + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-2}^2$ . Debido a su estructura,

rendimientos pasados grandes implican valores grandes de la varianza condicional. Esto significa que valores grandes de los rendimientos son seguidos, en promedio, de otro valor grande del rendimiento. Este hecho es similar a los agrupamientos de la volatilidad que se observan en las series de rendimientos financieros.

En la actualidad estos modelos son los más populares en la literatura financiera econométrica, tanto desde el punto de vista teórico como empírico, generando un gran número de artículos. Algunas revisiones se encuentran en Bollerslev et al. (1992), Bollerslev et al. (1994), Bera y Higgins (1995), Diebold y López (1995) y Palm (1997).

### 2.2.1 Modelos GARCH

La implementación de los modelos ARCH requería un número elevado de retardos de  $y_t^2$  en la ecuación de  $\sigma_t^2$ , haciendo estos modelos difíciles de tratar en la práctica. Por ello, Bollerslev (1986)<sup>2</sup> propuso un modelo más parsimonioso, capaz de representar la elevada persistencia observada a menudo en los cuadrados de las observaciones, el modelo ARCH Generalizado, o modelo GARCH.

Si  $y_t$  está generado por un modelo GARCH( $p, q$ ), entonces viene dado por (2.1) donde

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i y_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (2.2)$$

Si  $\beta_j = 0, \forall j \geq 1$ , se obtiene el modelo ARCH( $p$ ) como caso particular. El modelo GARCH más utilizado en la práctica es el GARCH(1,1), en el cual

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (2.3)$$

donde  $\omega$ ,  $\alpha$  y  $\beta$  son parámetros tales que  $\omega > 0$  y  $\alpha, \beta \geq 0$ . Las condiciones de positividad son necesarias para garantizar la existencia de la varianza condicional<sup>3</sup> y

<sup>2</sup>Taylor (1986) propuso el modelo GARCH(1,1) al mismo tiempo.

<sup>3</sup>Las condiciones de positividad para el modelo general GARCH( $p, q$ ) se encuentran en Nelson y Cao (1992).

$\omega$  tiene que ser estrictamente positiva para que el proceso  $y_t$  no degenera. Finalmente, si  $\alpha + \beta < 1$ , el proceso es estacionario en covarianza y la varianza marginal de  $y_t$  es finita y viene dada por

$$\sigma_y^2 = \frac{\omega}{1 - \alpha - \beta} \quad (2.4)$$

Nelson (1990) demuestra que, si  $E[\log(\beta + \alpha\varepsilon_t^2)] < 0^4$ , entonces  $y_t$  es estrictamente estacionario. Nótese que si  $\varepsilon_t$  es gaussiano, esta condición se satisface incluso si  $\alpha + \beta = 1$ . Por tanto, cuando  $\alpha + \beta = 1$  el proceso GARCH(1,1) es estrictamente estacionario, aunque la varianza marginal no sea finita. En la figura 2.4 se han representado las regiones de los parámetros  $\alpha$  y  $\beta$  que satisfacen las condiciones de estacionariedad en covarianza y estacionariedad estricta para distintas distribuciones de  $\varepsilon_t$ . En particular, se ha considerado la distribución gaussiana y dos distribuciones t de Student con 5 y 7 grados de libertad respectivamente. En dicho gráfico puede observarse que la región de estacionariedad en covarianza es la misma para las tres distribuciones. Sin embargo, la región de estacionariedad estricta depende de la distribución que se supone para las innovaciones. Nótese que dicha región aumenta a medida que los grados de libertad disminuyen<sup>5</sup>.

He y Teräsvirta (1999) y Ling y McAleer (2002b) prueban que la condición necesaria y suficiente para la existencia del momento de cuarto orden es  $\kappa_\varepsilon\alpha^2 + 2\alpha\beta + \beta^2 < 1$ , donde  $\kappa_\varepsilon$  es la curtosis de la innovación  $\varepsilon_t$ . Si esta condición se satisface, entonces la curtosis de  $y_t$  viene dada por

$$\kappa_y = \frac{E(y_t^4)}{[E(y_t^2)]^2} = \kappa_\varepsilon \frac{1 - \alpha^2 - \beta^2 - 2\alpha\beta}{1 - \kappa_\varepsilon\alpha^2 - 2\alpha\beta - \beta^2} \quad (2.5)$$

<sup>4</sup>Las condiciones para la estacionariedad estricta de los procesos generales GARCH( $p, q$ ) han sido derivadas por Bougerol y Picard (1992).

<sup>5</sup>Esto parece contradecir los resultados en Nelson (1990) y Kleibergen y van Dijk (1993). Sin embargo, nótese que en ambos trabajos se considera que la varianza de las innovaciones  $\varepsilon_t$  viene dada por la varianza de la distribución t de Student, mientras que aquí se supone que las innovaciones tienen varianza uno, lo que hace que el parámetro  $\alpha$  que se representa en la figura 2.4 sea mayor que el parámetro  $\alpha$  que se representa tanto en Nelson (1990) como en Kleibergen y van Dijk (1993).

que es mayor que 3 aunque  $\varepsilon_t$  sea gaussiano. Por lo tanto, la distribución marginal de  $y_t$  tiene colas pesadas. Además, puede verse fácilmente que todos sus momentos impares son cero.

La figura 2.4 muestra también, para las tres distribuciones de  $\varepsilon_t$  mencionadas anteriormente, las regiones donde los valores de los parámetros  $\alpha$  y  $\beta$  satisfacen la condición necesaria para la existencia de curtosis. Como se puede observar en este gráfico, independientemente de la distribución de  $\varepsilon_t$ , la región de estacionariedad estricta es la mayor de las tres regiones y contiene a las otras dos, lo que significa que si la curtosis del modelo es finita, el modelo es estacionario tanto en covarianza como de manera estricta. Se puede observar además que, a medida que disminuyen los grados de libertad de la distribución  $t$  de Student considerada, la región de los parámetros que garantiza que la curtosis del modelo existe, también disminuye. Nótese también que el valor máximo de  $\alpha$  que satisface la condición de existencia del momento de cuarto orden depende de la distribución. Por ejemplo, si la distribución de  $\varepsilon_t$  es una  $t$  de Student con 5 grados de libertad, entonces  $\alpha$  puede ser 0.3333 como máximo, si  $\beta = 0$ . Por otra parte, si  $\alpha = 0.15$ , para que la curtosis sea finita,  $\beta$  debe tomar un valor menor que 0.7555.

El modelo GARCH(1,1) puede escribirse, alternativamente, como un modelo ARMA(1,1) no gaussiano en los cuadrados, dado por

$$y_t^2 = \omega + (\alpha + \beta)y_{t-1}^2 + \nu_t - \beta\nu_{t-1} \quad (2.6)$$

donde  $\nu_t = y_t^2 - \sigma_t^2$ . La perturbación  $\nu_t$  es incorrelada, tiene media cero y varianza constante pero es condicionalmente heterocedástica. En la ecuación (2.6), es posible observar como la dinámica del proceso GARCH(1,1) aparece en la fac de las observaciones al cuadrado, que tiene la misma forma que la fac de un proceso ARMA(1,1).

Bollerslev (1988) deriva las autocorrelaciones de  $y_t^2$  que vienen dadas por

$$\begin{aligned}\rho_2(1) &= \frac{\alpha(1 - \alpha\beta - \beta^2)}{1 - 2\alpha\beta - \beta^2} \\ \rho_2(\tau) &= (\alpha + \beta)^{\tau-1} \rho_2(1), \tau > 1\end{aligned}\tag{2.7}$$

Nótese que la fac depende de los parámetros  $\alpha$  y  $\beta$  y que decae hacia cero geométricamente a la tasa  $\alpha + \beta$ . Por tanto, es común usar  $\alpha + \beta$  como una medida de la persistencia de las innovaciones en la volatilidad de los modelos GARCH(1,1). No obstante, debido a que es posible tener modelos ARCH estrictamente estacionarios que no tienen momentos finitos, la noción de persistencia de una innovación en la volatilidad en el marco de los procesos ARCH es más complicada que el correspondiente concepto de persistencia en la media para modelos lineales. Bauwens et al. (1999) discuten cuidadosamente la noción de persistencia en el proceso GARCH(1,1). La idea es estudiar si el efecto de una innovación en las varianzas condicionales futuras es persistente o si, por el contrario, desaparece con el tiempo. Para ello reescriben  $\sigma_t^2$  en (2.3) en función de los valores pasados de las innovaciones,  $\varepsilon_t$ , y de la varianza condicional inicial  $\sigma_0^2$ , como propuso Nelson (1990). Es fácil ver que

$$\sigma_t^2 = \omega + \sigma_{t-1}^2(\alpha\varepsilon_{t-1}^2 + \beta) = \sigma_0^2 \prod_{i=1}^t (\alpha\varepsilon_{t-i}^2 + \beta) + \omega \left[ 1 + \sum_{j=1}^{t-1} \prod_{i=1}^j (\alpha\varepsilon_{t-i}^2 + \beta) \right] \tag{2.8}$$

Nelson (1990) llama persistencia en  $L^1$  al impacto de la innovación inicial  $\varepsilon_0$  sobre la varianza condicional  $\sigma_t^2$ , esto es,

$$\lim_{t \rightarrow \infty} E \left[ \frac{\partial \sigma_t^2}{\partial \varepsilon_0^2} \right] \tag{2.9}$$

La razón para tomar esperanzas en (2.9) es que la derivada depende de todas las innovaciones  $\varepsilon_i$  entre los instantes 0 y  $t$ . Suponiendo que  $\sigma_0^2$  es independiente de la sucesión  $\varepsilon_t$ , se tiene que

$$E \left[ \frac{\partial \sigma_t^2}{\partial \varepsilon_0^2} \right] = E[\sigma_0^2] E \left[ \frac{\partial \prod_{i=1}^t (\alpha\varepsilon_{t-i}^2 + \beta)}{\partial \varepsilon_0^2} \right] =$$

$$E[\sigma_0^2]E\left[\prod_{i=1}^{t-1}(\alpha\varepsilon_{t-i}^2 + \beta)\right]E\left[\frac{\partial(\alpha\varepsilon_0^2 + \beta)}{\partial\varepsilon_0^2}\right] = E[\sigma_0^2][\alpha E(\varepsilon_t^2) + \beta]^{t-1}\alpha \quad (2.10)$$

De este modo, existirá persistencia (en el sentido de que el efecto de una innovación no desaparece nunca) si la expresión (2.10) converge a una cantidad finita y distinta de cero cuando  $t$  tiende a infinito. Y esto sucede siempre que  $\alpha E(\varepsilon_t^2) + \beta = 1$ . En este caso el impacto de la innovación será  $\alpha$ . Dado que la varianza de la perturbación  $\varepsilon_t$  es uno, tendremos persistencia cuando exista una raíz unitaria en el proceso GARCH(1,1), es decir, cuando  $\alpha + \beta = 1$ . Engle y Mustafa (1992) también usan  $\alpha[\alpha + \beta]^{t-1}$  como medida de persistencia.

La tabla 2.2 contiene los estimadores de Máxima Verosimilitud (MV) de los parámetros del modelo GARCH(1,1) con errores gaussianos<sup>6</sup> para las 12 series de rendimientos descritas en la tabla 2.1 junto con el logaritmo de la verosimilitud como medida de bondad de ajuste. Aquellas series que tenían alguna estructura en la media, en particular, los cinco índices bursátiles y dos de los siete tipos de cambio, han sido filtradas mediante modelos de medias móviles para recoger esa dependencia. Como puede observarse, todas las series tienen efectos ARCH significativos y valores de  $\hat{\alpha} + \hat{\beta}$  muy próximos a uno. El diagnóstico de los modelos se basa en las observaciones estandarizadas, definidas como  $\hat{\varepsilon}_t = y_t/\hat{\sigma}_t$ , donde  $\hat{\sigma}_t$  se obtiene sustituyendo los parámetros por sus estimadores en la expresión correspondiente de la varianza condicional. La figura 2.5 representa las volatilidades estimadas,  $\hat{\sigma}_t$ , correspondientes a las 12 series. La tabla 2.2, muestra también algunos momentos muestrales de  $\hat{\varepsilon}_t$ . Nótese que las observaciones estandarizadas son asimétricas y tienen colas pesadas. Esto se ve más claramente en la figura 2.6 donde se han representado estimaciones kernel de las densidades de  $\hat{\varepsilon}_t$ , junto con la densidad gaussiana. Por lo tanto, si  $\sigma_t$  se modeliza mediante un modelo GARCH(1,1), el supuesto de gaussianidad para  $\varepsilon_t$  no parece adecuado. Sin embargo, las autocorrelaciones de  $\hat{\varepsilon}_t^2$  ya no son significativas. En consecuencia, para estas series, el

<sup>6</sup>La estimación se ha llevado a cabo con el programa EViews, versión 3.1.



modelo GARCH(1,1) gaussiano es capaz de representar adecuadamente la dinámica de los cuadrados pero no puede explicar ni las asimetrías ni el exceso de curtosis presente en los datos. Esto puede deberse, como se verá en detalle en los siguientes capítulos, a un supuesto no adecuado de la distribución de  $\varepsilon_t$  y/o sobre la especificación de  $\sigma_t$  y/o a la presencia de observaciones atípicas.

El supuesto de gaussianidad para las innovaciones  $\varepsilon_t$  ha sido relajado por varios autores. Por ejemplo, Bollerslev (1987) sugiere una distribución t de Student, Jorion (1988) propone usar una mixtura de distribuciones Normal y Poisson. El uso de la mixtura de Normal y Lognormal es propuesta por Hsieh (1989) y la distribución GED (Generalized error distribution) por Nelson (1991). Bollerslev et al. (1994) usan además la distribución t Generalizada que incluye a las distribuciones t de Student y GED como casos particulares. Granger y Ding (1995) y González-Rivera (1998) también consideran el uso de la distribución de Laplace. Finalmente, Jiang et al. (2001) suponen que la distribución de las innovaciones  $\varepsilon_t$  es desconocida y usan métodos semiparamétricos para estimar los parámetros de los modelos ARCH.

La tabla 2.3 contiene los estimadores MV de los parámetros del modelo GARCH(1,1) con errores t de Student. Comparando estas estimaciones con las de los modelos GARCH(1,1) gaussianos, se puede observar que, con la excepción del índice de Bombay, no hay grandes diferencias entre las estimaciones de los parámetros. Sin embargo, en los rendimientos del índice de Bombay,  $\hat{\alpha}$  cambia dramáticamente cuando se supone una distribución condicional t de Student. Bajo el supuesto de gaussianidad, el parámetro  $\alpha$  se estima como 0.0434 con un error estándar de 0.006 y cuando se supone una t de Student pasa a estimarse como 0.1317 con un error estándar de 0.034. Este fenómeno puede ser debido a la presencia de observaciones atípicas, como se verá en el Capítulo 4. En cuanto a las observaciones estandarizadas, nótese que ya no tienen exceso de curtosis con respecto a la distribución t de Student. Por tanto, los modelos GARCH(1,1)-t parecen capaces de explicar tanto la dinámica de los cuadrados como la elevada cur-

tosis que se observa en las series. Sin embargo, en cinco de las doce series analizadas, los parámetros estimados violan la condición de existencia del momento de cuarto orden, por lo que aparentemente, no existiría la curtosis, algo que parece muy restrictivo cuando se mira a las series reales. En el Capítulo 3 se verá una posible explicación a este fenómeno.

Las tablas 2.4, 2.5 y 2.6 muestran los valores de la media, varianza, curtosis y autocorrelaciones de los cuadrados que implican los modelos GARCH(1,1) estimados para cada una de las series, junto con sus valores muestrales. Nótese que, con la excepción del índice de Amsterdam, cuando los errores se suponen gaussianos, la curtosis teórica que implica el modelo es menor que la curtosis muestral. Por el contrario, si se supone una distribución *t* de Student, en los casos en que los parámetros satisfacen la condición de existencia del momento de cuarto orden, la curtosis que implica el modelo es siempre mayor que la muestral. Por tanto, parece que el modelo GARCH(1,1) no es capaz de explicar bien la curtosis de estas series ya que cuando se supone gaussianidad condicional la curtosis que implica el modelo es demasiado baja y cuando se supone una distribución leptocúrtica la curtosis que implica el modelo es demasiado alta, cuando existe.

Dado que en muchas aplicaciones empíricas y como muestra la tabla 2.2, es común obtener estimaciones de  $\alpha$  y  $\beta$  tales que  $\hat{\alpha} + \hat{\beta} \simeq 1$ , Engle y Bollerslev (1986) propusieron el modelo GARCH Integrado (IGARCH) dado por (2.3) con  $\alpha + \beta = 1$ . Es importante recordar que aunque la varianza marginal del proceso no sea finita, el proceso puede ser estrictamente estacionario.

Alternativamente, considerando que, como se ha visto con un ejemplo en la figura 2.3, las autocorrelaciones de los rendimientos en valor absoluto suelen ser mayores que las de los cuadrados de los rendimientos, varios autores han propuesto modelizar la desviación típica condicional en lugar de la varianza condicional<sup>7</sup>. Los primeros

---

<sup>7</sup>No obstante, He y Teräsvirta (1999) sugieren que el *efecto Taylor* podría ser debido a sesgos en

modelos para el valor absoluto de los rendimientos se remontan a Taylor (1986) y Schwert (1989) quienes propusieron el modelo AVGARCH (Absolute Value GARCH) dado por

$$\sigma_t = \omega + \alpha |y_{t-1}| + \beta \sigma_{t-1} \quad (2.11)$$

Nelson y Foster (1994) demuestran que el modelo anterior es un filtro más eficiente de la varianza condicional en presencia de innovaciones con distribuciones leptocúrticas que las especificaciones basadas en  $\sigma_t^2$ . He y Teräsvirta (1999) derivan la fac de los cuadrados en el modelo AVGARCH que es radicalmente diferente de la del modelo GARCH. Aunque la varianza marginal, curtosis y fac de las observaciones al cuadrado del modelo AVGARCH son inmediatas a partir de los resultados de He y Teräsvirta (1999), sus expresiones son muy complicadas y por lo tanto no se reproducen aquí.

Los modelos GARCH tienen muchos atractivos, por lo que, como ya se ha comentado, su implementación empírica es muy extensa. No obstante, estos modelos tienen también algunas limitaciones señaladas originalmente por Nelson (1991) y, entre las que Tsay (2002) destaca:

1. Suponen que los rendimientos positivos y negativos tienen el mismo efecto en la volatilidad ya que ésta depende de los valores previos al cuadrado. Sin embargo, en la práctica puede existir una respuesta asimétrica de la volatilidad ante movimientos positivos y negativos de los precios. Esto es conocido como *efecto apalancamiento* y fue descrito originalmente por Black (1976). Esta asimetría ha sido también documentada por Glosten et al. (1993), Schwert (1989), Nelson (1991), Campbell y Hentschel (1992), Engle y Ng (1993), Sentana (1995) y Shephard (1996) entre otros.
2. Imponen restricciones sobre los parámetros que van aumentando a medida que se quiere garantizar la existencia de momentos de mayor orden y que se permiten

---

la fac muestral de los cuadrados.

colas más anchas para la distribución condicional. Además, se ha observado en algunas aplicaciones empíricas, que las restricciones de no negatividad sobre los parámetros no se satisfacen.

3. Tienden a sobrepredecir la volatilidad porque dependen directamente de los valores grandes aislados de la serie de rendimientos. Esto hace que las estimaciones de la volatilidad sean muy poco suaves.
4. El concepto de persistencia en estos modelos no es claro.

### 2.2.2 Modelos EGARCH

Para solucionar algunas de las limitaciones de los modelos GARCH, mencionadas anteriormente, Nelson (1991) propone el modelo EGARCH (Exponential GARCH). El modelo más simple, EGARCH(1,1), viene dado por  $y_t$ , en (2.1) con

$$\log \sigma_t^2 = \omega + \beta \log \sigma_{t-1}^2 + \alpha \left[ \left| \frac{y_{t-1}}{\sigma_{t-1}} \right| - E \left| \frac{y_t}{\sigma_t} \right| \right] + \gamma \frac{y_{t-1}}{\sigma_{t-1}} \quad (2.12)$$

donde  $E \left| \frac{y_{t-1}}{\sigma_{t-1}} \right| = \sqrt{\frac{2}{\pi}}$ , si  $\varepsilon_t$  es gaussiano. Nótese que en (2.12) no hay necesidad de restringir los parámetros para garantizar que la varianza condicional sea positiva ya que el modelo está formulado para el proceso  $\log \sigma_t^2$ . La condición de estacionariedad es  $|\beta| < 1$ . Este modelo implica una respuesta asimétrica de la volatilidad ante rendimientos positivos y negativos debida al término  $\gamma \frac{y_{t-1}}{\sigma_{t-1}}$ . Si, por ejemplo, el coeficiente  $\gamma$  es negativo, esto implica que rendimientos positivos generan menos volatilidad que rendimientos negativos.

La figura 2.7 contiene la Curva NIC (News Impact Curve) propuesta por Engle y Ng (1993) para medir el impacto de las innovaciones sobre la volatilidad. Esta curva mide la relación entre  $y_{t-1}$  y  $\sigma_t^2$  cuando toda la información hasta el instante  $t-2$ , incluido, se mantiene constante y todos los retardos de las varianzas condicionales están

evaluados en la varianza marginal. Para el modelo GARCH(1,1), la curva NIC tiene la siguiente expresión

$$\sigma_t^2 = A + \alpha y_{t-1}^2 \quad (2.13)$$

donde  $A = \omega + \beta\sigma_y^2$ . Como se puede observar, la curva NIC para el modelo GARCH es simétrica, lo que significa que innovaciones positivas tienen el mismo efecto sobre la volatilidad que innovaciones negativas siempre y cuando ambas sean iguales en valor absoluto. Al igual que en el modelo GARCH(1,1), la curva NIC del modelo AVGARCH, también representada en la figura 2.7, es simétrica, de modo que este modelo tampoco permite asimetrías en la volatilidad. Sin embargo, para el modelo EGARCH(1,1) gaussiano, la curva NIC viene dada por

$$\sigma_t^2 = \begin{cases} A \exp\left(\frac{\gamma+\alpha}{\sigma_y} y_{t-1}\right) & \text{si } y_{t-1} > 0 \\ A \exp\left(\frac{\gamma-\alpha}{\sigma_y} y_{t-1}\right) & \text{si } y_{t-1} < 0 \end{cases} \quad (2.14)$$

donde  $A = \sigma_y^{2\beta} \exp(\omega - \alpha\sqrt{(2/\pi)})$ . Considerando  $\omega > 0$ ,  $0 < \alpha < 1$ ,  $0 < \beta < 1$  y, muy importante,  $\gamma < 0$ , como se ha hecho en la figura 2.7, las innovaciones negativas tienen un efecto sobre la volatilidad mayor que las innovaciones positivas.

Expresiones para la varianza marginal, curtosis y fac de las observaciones al cuadrado del proceso EGARCH en (2.12) se encuentran en He et al. (1999) y vienen dadas por

$$\sigma_y^2 = \exp\left(\frac{\omega}{1-\beta}\right) \prod_{i=1}^{\infty} E(\exp(\beta^{i-1}g)) \quad (2.15)$$

$$\kappa_y = 3 \prod_{i=1}^{\infty} \frac{E(\exp(2\beta^{i-1}g))}{[E(\exp(\beta^{i-1}g))]^2} \quad (2.16)$$

y

$$\rho_2(\tau) = \frac{E(\varepsilon_{t-1}^2 \exp(\beta^{\tau-1}g)) P_1 P_2 - P_3}{3P_4 - P_3} \quad (2.17)$$

donde, si  $\varepsilon_t$  es gaussiano,  $g = g\left(\frac{y_{t-1}}{\sigma_{t-1}}\right) = \alpha\left(\left|\frac{y_{t-1}}{\sigma_{t-1}}\right| - \sqrt{2/\pi}\right) + \gamma\frac{y_{t-1}}{\sigma_{t-1}}$ ,  $P_1 = \prod_{i=1}^{\tau} E(\exp(\beta^{i-1}g))$ ,  $P_2 = \prod_{i=1}^{\infty} E(\exp((1+\beta^{\tau})\beta^{i-1}g))$ ,  $P_3 = \prod_{i=1}^{\infty} [E(\exp(2\beta^{i-1}g))]^2$  and  $P_4 = \prod_{i=1}^{\infty} E(\exp(2\beta^{i-1}g))$ .

La figura 2.8 representa la fac de los cuadrados del modelo EGARCH(1,1) para distintos valores de los parámetros. Una conclusión inmediata de este gráfico es que la tasa a la que decae dicha fac no es  $\beta$ , puesto que para el mismo valor de  $\beta$ , y para un valor fijado de  $\alpha$ , la fac decae a distinta tasa dependiendo del valor del parámetro de asimetría,  $\gamma$ . Análogamente, para el mismo valor de  $\beta$  y fijando un valor de  $\gamma$ , la fac decae a distinta tasa dependiendo del valor de  $\alpha$ .

Nelson (1991), Kearns y Pagan (1993), Poon y Taylor (1992), Zakoian (1994) y Chong et al. (1999) entre otros, han ajustado modelos EGARCH a distintas series temporales financieras reales. No obstante, a pesar de las ventajas que ofrecen los modelos EGARCH frente a los GARCH, dichos modelos no han sido muy utilizados en la práctica, posiblemente debido a la falta de resultados asintóticos acerca de los estimadores de Cuasi-Máxima Verosimilitud y a que los modelos EGARCH tienden a sobreestimar el efecto de las observaciones extremas en la volatilidad; véase, por ejemplo, Hafner y Herwartz (2000).

La tabla 2.7 contiene los estimadores MV de los parámetros del modelo EGARCH(1,1) con errores gaussianos<sup>8</sup> para las 12 series. Obsérvese que el coeficiente de persistencia de la volatilidad,  $\beta$ , se estima muy próxima a uno y el parámetro de asimetría,  $\gamma$ , es significativo para todas las series analizadas, excepto para los tipos de cambio US-CAN y US-ESP. Además, nótese que, excepto para el tipo de cambio US-SWK, dicho parámetro se estima negativo, lo que implica que los rendimientos negativos tienen un efecto sobre la volatilidad mayor que los positivos. Finalmente, puede observarse que las observaciones estandarizadas,  $\hat{\varepsilon}_t$ , tienen exceso de curtosis y, para ocho de las doce series, son también asimétricas. Por lo tanto, como en el caso de los modelos GARCH(1,1), los modelos EGARCH(1,1) gaussianos son capaces de representar la dinámica de los cuadrados pero no pueden explicar ni el exceso de curtosis ni todas las asimetrías presentes en los datos. En las tablas 2.4, 2.5 y 2.6 puede observarse que los

---

<sup>8</sup>La estimación se ha llevado a cabo con el programa EViews, versión 3.1.

valores de la curtosis y las autocorrelaciones de los cuadrados que implican los modelos EGARCH(1,1) estimados son muy parecidos a los valores que implican los modelos GARCH(1,1), aunque en el caso de la varianza, los valores que implican los modelos EGARCH(1,1) están muy alejados de los muestrales.

### 2.2.3 Otros modelos ARCH para varianzas condicionales asimétricas

Desde la propuesta original de Nelson (1991), se han propuesto varios modelos alternativos para representar la respuesta asimétrica de la volatilidad. Entre los modelos asimétricos más populares se encuentra el modelo A-PARCH (Asymmetric-Power ARCH), propuesto por Ding et al. (1993), que es capaz de unificar a siete modelos tipo ARCH. En particular, los modelos GARCH, AVGARCH, GJR-GARCH de Glosten et al. (1993), TARCH (Threshold-ARCH) de Zakoian (1994), NARCH de Higgins y Bera (1992) y log-ARCH de Geweke (1986) y de Pantula (1986) son casos particulares del modelo A-PARCH.

La varianza condicional en el modelo A-PARCH más simple viene dada por

$$\sigma_t^\delta = \omega + \alpha(|y_{t-1}| - \gamma y_{t-1})^\delta + \beta \sigma_{t-1}^\delta \quad (2.18)$$

Ding et al. (1993) y Granger y Ding (1995) señalan que el parámetro  $\delta$  sirve como una transformación Box-Cox de  $\sigma_t$  y que es necesario para captar adecuadamente la dinámica de la volatilidad. El parámetro  $\gamma$  modeliza la respuesta asimétrica de  $\sigma_t$  ante innovaciones positivas y negativas. Las propiedades estadísticas del modelo A-PARCH han sido tratadas por He y Teräsvirta (1997) y Fornani y Mele (1997). Este modelo ha sido ajustado a rendimientos de varios mercados financieros por Brooks et al. (2000) y Paolella (2000).

Recientemente, han aparecido varios modelos alternativos que permiten obtener

como casos particulares la mayoría de los modelos tipo ARCH descritos anteriormente. En primer lugar, Hentschel (1995) define una familia paramétrica de modelos GARCH que incluye a los modelos EGARCH y A-PARCH pero no al modelo GQARCH de Sentana (1995). León y Mora (1999) ajustan el modelo propuesto por Hentschel (1995) a rendimientos diarios del índice de la Bolsa de Madrid, IBEX-35. Para dicha serie, los modelos que representan la desviación típica condicional se ajustan mejor que los que modelizan las varianzas condicionales. Además, la verosimilitud de los modelos basados en distribuciones condicionales leptocúrticas es mayor que cuando la distribución condicional se supone Normal. Finalmente, muestran que el comportamiento asimétrico de los rendimientos del IBEX-35 es estadísticamente significativo. La respuesta asimétrica de la volatilidad y la distribución condicional leptocúrtica para los mismos rendimientos se menciona también en Blanco (2000).

Posteriormente, Duan (1997) introduce el modelo GARCH Aumentado que es lo bastante general como para unificar a los principales modelos tipo ARCH propuestos en la literatura. Por ejemplo, los modelos GARCH, EGARCH y GJR-GARCH son, entre otros, casos particulares del modelo GARCH Aumentado. Loudon et al. (2000), ajustan dicho modelo a los rendimientos diarios del índice *FT All Share* del Reino Unido, observados desde 1971 a 1997. Concluyen que la volatilidad de dicha serie exhibe un alto grado de asimetría y de persistencia (en el sentido de que el efecto de una innovación sobre la volatilidad acaba desapareciendo aunque lo hace muy lentamente). Sin embargo, los residuos se caracterizan por tener coeficientes de asimetría y curtosis significativos. Por lo tanto, concluyen que los modelos tipo ARCH con funciones de densidad condicional normales son capaces de captar parte, aunque no toda la asimetría y exceso de curtosis observado. La misma conclusión se encuentra en McCurdy y Morgan (1987), Milhoj (1987), Hsieh (1989) y Baillie y Bollerslev (1989).

Finalmente, He y Teräsvirta (1999) proponen un marco unificador para considerar las propiedades estadísticas de muchos modelos GARCH(1,1), simétricos y asimétricos,



y sin necesidad de hacer supuestos particulares sobre la distribución de  $\varepsilon_t$ . Entre los modelos considerados se encuentran los siguientes: GARCH(1,1), AVGARCH(1,1), GJR-GARCH(1,1), GARCH(1,1) no lineal, que es un caso particular del modelo APARCH, “volatility switching” GARCH(1,1) de Fornani y Mele (1997), TGARCH(1,1), “fourth-order nonlinear generalized moving-average conditional heteroscedasticity” de Yang y Bewley (1995) y GQARCH(1,1). No incluyen ni el modelo EGARCH ni el APARCH. Para su familia de modelos GARCH(1,1), derivan una condición general para la existencia de cualquier momento de orden entero, tanto para las observaciones como para el valor absoluto de las observaciones, junto con la fac de los cuadrados de las observaciones y de su valor absoluto. Ling y McAleer (2002b) consideran esta misma familia de modelos GARCH(1,1) para la cual establecen una condición suficiente para la estacionariedad estricta y la condición necesaria y suficiente para la existencia de sus momentos.

En la actualidad hay un gran interés en las propiedades teóricas de esta familia de modelos lo que ha dado lugar a varios artículos que aportan condiciones necesarias y suficientes para la estacionariedad en covarianza y estacionariedad estricta así como para la existencia de momentos de cualquier orden. Véase también Li et al. (2001) y Ling y McAleer (2002a).

## 2.3 Modelos de Volatilidad Estocástica

Los modelos ARSV (Autoregressive Stochastic Volatility), propuestos originalmente por Taylor (1986) y popularizados por Harvey et al. (1994), especifican la volatilidad,  $\sigma_t$ , como una variable latente que, en el caso más sencillo, sigue un proceso autorregresivo una vez se ha transformado tomando logaritmos.



El caso más simple es el modelo ARSV(1)<sup>9</sup> que viene dado por

$$\begin{aligned} y_t &= \sigma_* \varepsilon_t \sigma_t \\ \log \sigma_t^2 &= \phi \log \sigma_{t-1}^2 + \eta_t \end{aligned} \quad (2.19)$$

donde se asume que  $\varepsilon_t$  y  $\eta_t$  son procesos i.i.d., mutuamente independientes, con media cero y varianzas uno y  $\sigma_\eta^2$  respectivamente. Se diferencian, por tanto, dos perturbaciones, una,  $\varepsilon_t$ , que afecta al nivel de la serie y otra,  $\eta_t$ , que afecta a la volatilidad. De este modo, la volatilidad tiene un componente predecible que depende del pasado y un componente inesperado. Además,  $\eta_t$  se asume Normal. El parámetro  $\sigma_*$  es un factor de escala que evita incluir un término constante en la ecuación de  $\log \sigma_t^2$  y la restricción  $|\phi| < 1$  garantiza la estacionariedad de  $y_t$ . A pesar de que el supuesto de Normalidad para  $\eta_t$  pueda parecer *ad hoc*, se ha observado que la distribución del proceso log-volatilidad en series financieras reales puede aproximarse adecuadamente por una distribución Normal; ver, por ejemplo, Andersen, Bollerslev, Diebold, y Ebens (2001) y Andersen, Bollerslev, Diebold, y Labys (2001). Nótese que  $\sigma_\eta^2$  es la varianza de la perturbación de la volatilidad. Cuando  $\sigma_\eta^2$  es cero, el modelo (2.19) no está identificado. So et al. (1999) presentan una interpretación muy interesante del modelo ARSV(1) mediante la descomposición de la volatilidad total de  $y_t$  en dos términos: la volatilidad de un día típico, y la volatilidad debida a las fluctuaciones por las llegadas de información al mercado. Taylor (1994), Ghysels et al. (1996) y Shephard (1996) contienen revisiones de la literatura sobre los modelos ARSV.

En general, la distribución de  $y_t$  condicionada a las observaciones pasadas no tiene una forma conocida. Sin embargo, usando las propiedades de la distribución log-normal, se pueden obtener distintos momentos de la distribución marginal. Así, la varianza de  $y_t$  viene dada por

$$\sigma_y^2 = \sigma_*^2 \exp(0.5\sigma_\eta^2) \quad (2.20)$$

---

<sup>9</sup>La especificación del proceso log-volatilidad como un proceso ARMA( $p, q$ ) ha sido considerado, por ejemplo, por Hwang y Satchell (2000).

donde  $\sigma_h^2 = \sigma_\eta^2 / (1 - \phi^2)$ . La curtosis de  $y_t$  es

$$\kappa_y = \kappa_\varepsilon \exp(\sigma_h^2) \quad (2.21)$$

donde  $\kappa_\varepsilon$  es la curtosis de  $\varepsilon_t$ . Aunque  $\varepsilon_t$  sea gaussiano,  $\kappa_y$  es mayor que 3 y, por tanto, el modelo ARSV(1) también es capaz de generar series con colas pesadas. Nótese que, en el modelo ARSV(1), el parámetro  $\sigma_\eta^2$  controla el grado de curtosis, independientemente de la persistencia de las innovaciones en la volatilidad, medida por  $\phi$ . Podría haber exceso de curtosis incluso cuando  $\phi = 0$ . Como se verá en el capítulo siguiente, esta propiedad hace a los modelos ARSV(1) más flexibles que los modelos GARCH(1,1) a la hora de representar los hechos empíricos usualmente observados en series reales. Además, en el modelo ARSV(1), la condición para la existencia del momento de cuarto orden es la condición de estacionariedad, es decir,  $|\phi| < 1$ . Esta propiedad es una ventaja frente al modelo GARCH(1,1), que puede ser estacionario y no tener momento de cuarto orden finito, como se veía en el gráfico 2.4.

Se puede ver fácilmente que  $y_t$  es un proceso incorrelado aunque no independiente cuyos momentos impares son cero. Las autocorrelaciones de las observaciones al cuadrado han sido derivadas por Taylor (1986) y vienen dadas por

$$\rho_2(\tau) = \frac{\exp(\sigma_h^2 \phi^\tau) - 1}{\kappa_\varepsilon \exp(\sigma_h^2) - 1} \quad (2.22)$$

Taylor (1986) menciona que cuando  $\sigma_h^2$  es pequeña y/o las autocorrelaciones de  $\log \sigma_t^2$  son próximas a uno, la fac de las observaciones al cuadrado se puede aproximar mediante la fac de  $\log \sigma_t^2$  multiplicada por un factor de proporcionalidad, esto es,

$$\rho_2(\tau) \simeq \frac{\exp(\sigma_h^2) - 1}{\kappa_\varepsilon \exp(\sigma_h^2) - 1} \phi^\tau \quad (2.23)$$

Consecuentemente, se ha argumentado con frecuencia que la fac de las observaciones al cuadrado en el modelo ARSV(1) es similar a la de un proceso ARMA(1,1), caracterizada por un decaimiento geométrico determinado por el parámetro  $\phi$ ; ver, por

ejemplo, Taylor (1994) y Shephard (1996). Sin embargo, dicha aproximación puede no ser adecuada para determinados valores de los parámetros  $\phi$  y  $\sigma_\eta^2$  de interés en aplicaciones empíricas. Como ilustración, en la figura 2.9 se ha representado la fac para las observaciones al cuadrado del modelo ARSV(1) junto con la aproximación de Taylor para algunos valores seleccionados de los parámetros. En primer lugar, nótese que la aproximación siempre está por encima de la verdadera fac. Además, para la mayoría de los casos considerados, el decaimiento de la fac no es geométrico sino que es más rápido en los primeros retardos. Este hecho se puede observar más claramente en la figura 2.10 que representa, para los mismos valores de los parámetros considerados anteriormente, el cociente entre dos valores consecutivos de la fac, que viene dado por la siguiente expresión

$$\frac{\rho_2(\tau)}{\rho_2(\tau-1)} = \frac{\exp(\phi^\tau \sigma_h^2) - 1}{\exp(\phi^{(\tau-1)} \sigma_h^2) - 1} = \phi \frac{1 + 0.5\sigma_h^2 \phi^\tau + 0.17\sigma_h^4 \phi^{2\tau} + \dots}{1 + 0.5\sigma_h^2 \phi^{(\tau-1)} + 0.17\sigma_h^4 \phi^{2(\tau-1)} + \dots} \quad (2.24)$$

Nótese que, a medida que los retardos considerados son mayores, se tiende al decaimiento geométrico a la tasa  $\phi$ . Por su aparente similitud con la fac de los cuadrados de un modelo GARCH(1,1), en la que  $\alpha + \beta$  da una idea de la persistencia de las innovaciones en la volatilidad, en los modelos ARSV(1) la persistencia de las innovaciones en la volatilidad se ha medido, habitualmente, por  $\phi$ .

En la tabla 2.8 se muestran las estimaciones de los parámetros de los modelos ARSV(1) para las doce series anteriores. El método de estimación utilizado es el método propuesto por Sandmann y Koopman (1998)<sup>10</sup> que consiste en maximizar la función de verosimilitud mediante Monte Carlo<sup>11</sup>. El modelo ARSV(1) puede expresarse como un modelo en el espacio de los estados con innovaciones no gaussianas una vez se toma

<sup>10</sup>Este método se ha implementado usando el lenguaje de programación *Ox 3.1* de Doornik (2001) y usando la librería *SsfPack 2.3* de Koopman y Doornik (1999). El programa empleado para la estimación se llama *sv - mcl.ox* y se ha obtenido de la página web <http://www.econ.vu.nl/koopman/sv/>. Más información en <http://www.nuff.ox.ac.uk/Users/Doornik/index.html> y <http://www.ssfpack.com>.

<sup>11</sup>Para las series US-SSF y US UKP este método no converge, por lo que, para dichas series, los parámetros han sido estimados mediante el método de Cuasi Máxima Verosimilitud propuesto in

el logaritmo de las observaciones al cuadrado, esto es,

$$\begin{aligned}\log y_t^2 &= \mu + \log \sigma_t^2 + \xi_t \\ \log \sigma_t^2 &= \phi \log \sigma_{t-1}^2 + \eta_t\end{aligned}\tag{2.25}$$

donde  $\mu = \log \sigma_\star^2 + E(\log \varepsilon_t^2)$  y  $\xi_t = \log \varepsilon_t^2 - E(\log \varepsilon_t^2)$ . Nótese que cuando  $\varepsilon_t$  es Normal,  $\xi_t$  sigue una distribución  $\log(\chi_{(1)}^2)$  cuya media y varianza son  $-1.27$  y  $\pi^2/2$  respectivamente. El método de estimación propuesto por Sandmann y Koopman (1998) consiste en aproximar la función de verosimilitud mediante la suma de una parte gaussiana, construida mediante el filtro de Kalman y una función resto cuya esperanza puede ser evaluada por simulación.

Un resultado interesante que se observa comparando las estimaciones ARSV(1) en la tabla 2.8 y las estimaciones GARCH(1,1) en la tabla 2.2 es que, para la mayoría de las series, el parámetro  $\phi$ , asociado con la persistencia en el modelo ARSV(1), se estima menor que el correspondiente  $\alpha + \beta$  en el modelo GARCH(1,1). Este hecho ha sido ya mencionado en la literatura por varios autores; ver, por ejemplo, Shephard (1996) y Taylor (1994). En la tabla 2.8, se puede observar además que las observaciones estandarizadas, para la mayoría de las series, ya no tienen asimetrías ni exceso de curtosis y el supuesto de Normalidad parece adecuado. La figura 2.6 contiene estimadores kernel de la densidad de las observaciones estandarizadas con modelos GARCH(1,1) y ARSV(1) junto con la densidad Normal para las doce series consideradas. Como se puede ver en esta figura, la distribución de las observaciones estandarizadas con el modelo ARSV(1) es más próxima a la distribución Normal, que cuando el modelo considerado es el GARCH(1,1). En el siguiente capítulo abordaremos estos dos temas con más detalle.

Por otra parte, el exceso de curtosis y las asimetrías que se observan en las observaciones estandarizadas de algunas series, tanto cuando se estima un modelo GARCH(1,1) dependientemente por Harvey et al. (1994) y Nelson (1988), basado en maximizar la función de verosimilitud gaussiana del logaritmo de las observaciones al cuadrado.

como un modelo ARSV(1), puede deberse a la presencia de observaciones atípicas, como ya se ha comentado anteriormente. La presencia de observaciones atípicas en series condicionalmente heterocedásticas será considerado en detalle en el Capítulo 4.

A pesar de que el modelo ARSV(1) parece explicar las series consideradas, de un modo más adecuado que el modelo GARCH(1,1), es importante señalar que, en la tabla 2.8, el estadístico de Box-Ljung es significativo, al 95% de confianza, para la mayoría de las observaciones estandarizadas, lo que indica que hay estructura en los residuos del modelo. Sin embargo, dado que las correlaciones en los cuadrados son pequeñas y negativas y son significativas solamente para las series estimadas por medio del método propuesto por Sandmann y Koopman (1998), parece éste un problema del modo en que se estima la volatilidad más que del propio modelo ARSV(1). Podría deberse al filtro que utilizan los autores para obtener los valores suavizados  $\hat{\sigma}_t$ .

Las tablas 2.4, 2.5 y 2.6 muestran los valores de la media, varianza, curtosis y autocorrelaciones de los cuadrados que implican los modelos ARSV(1) estimados para cada una de las series junto con sus valores muestrales. Nótese que, al igual que en el caso de los modelos GARCH(1,1), para la mayoría de las series, la curtosis teórica que implica el modelo ARSV(1) es menor que la curtosis muestral. Sin embargo, la curtosis de los modelos ARSV(1) está, en cualquier caso, más cerca de la muestral.

Finalmente, la figura 2.5 contiene estimaciones de la volatilidad, obtenidas mediante los modelos ARSV(1). Nótese que la especificación ARSV(1) para la volatilidad permite obtener estimadores suavizados de  $\sigma_t$  usando toda la muestra,  $(y_1, \dots, y_T)$ ; véase, por ejemplo, Harvey et al. (1994). Como se puede ver en la figura 2.5, comparando las estimaciones de la volatilidad con las obtenidas mediante los modelos GARCH(1,1), ambas son muy similares aunque, en general, la especificación ARSV(1) produce estimadores de la volatilidad más suaves que la especificación GARCH(1,1). En los modelos GARCH(1,1) las observaciones en el instante  $t$  e instantes posteriores no modifican la estimación de  $\sigma_t$ , puesto que  $\sigma_t$  es observable en el momento  $t - 1$ .

## 2.4 Otros modelos alternativos

Aunque los modelos descritos en las secciones anteriores son los más conocidos y utilizados en la práctica, la literatura sobre modelos para representar la evolución dinámica de la volatilidad,  $\sigma_t$ , es muy extensa. Hay otros muchos modelos que pueden ser alternativas interesantes a los modelos anteriores; como, por ejemplo, los modelos CHARMA (Conditional Heteroscedastic Autoregressive Moving Average), que fueron introducidos por Tsay (1987). El modelo CHARMA más simple viene dado por

$$\begin{aligned}\phi(L)y_t &= \theta(L)a_t \\ \delta_t(L)a_t &= \eta_t\end{aligned}\tag{2.26}$$

donde  $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$  y  $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q$  son polinomios en  $L$ , de grados  $p$  y  $q$  respectivamente, y de coeficientes constantes.  $L$  es el operador de retardos tal que  $L^j y_t = y_{t-j}$ ,  $\eta_t$  es un proceso ruido blanco gaussiano con varianza  $\sigma_\eta^2$  y  $\delta_t(L) = 1 - \delta_{1t} L - \delta_{2t} L^2 - \dots - \delta_{rt} L^r$  es un polinomio en  $L$ , de grado  $r$ , con coeficientes aleatorios. El vector de coeficientes aleatorios  $\delta_t = (\delta_{1t}, \delta_{2t}, \dots, \delta_{rt})'$  forma una secuencia de vectores aleatorios i.i.d. de media cero y matriz de varianzas y covarianzas definida positiva,  $\Sigma$ . Además,  $\delta_t$  es independiente de  $\eta_t$ . La presencia de coeficientes aleatorios genera heterocedasticidad condicional en los modelos CHARMA. Si, por ejemplo,  $p = q = 0$  y  $r = 1$ , entonces la serie  $y_t$  viene dada por

$$y_t = \delta_{1t} y_{t-1} + \eta_t$$

Es sencillo probar que el proceso  $y_t$  es incorrelado, condicionalmente heterocedástico y con una distribución marginal leptocúrtica. Tsay (1987) considera una aplicación del modelo CHARMA utilizando la estructura de heterocedasticidad condicional como una manera alternativa de tratar observaciones atípicas en series temporales.

Otros modelos alternativos a los GARCH y ARSV son los modelos All-Pass, propuestos por Breidt et al. (2001). Estos modelos son modelos ARMA en los cuales

todas las raíces del polinomio autorregresivo son recíprocas de las del polinomio de medias móviles y viceversa. Este tipo de modelos generan series incorreladas pero no independientes en el caso de innovaciones no gaussianas.

Breidt et al. (2001) aplican esta metodología a rendimientos de tipos de cambio mostrando que los modelos All-Pass lineales pueden captar comportamientos no lineales asociados a menudo con los modelos GARCH y ARSV. El modelo All-Pass más simple es el de orden uno dado por

$$(1 - \phi L)y_t = (1 - \phi^{-1}L)a_t$$

donde  $a_t$  se supone i.i.d. de media cero y varianza  $\sigma^2$ . Se puede ver fácilmente que  $y_t$  es una secuencia incorrelada y, en el caso de que  $a_t$  sea Normal, la distribución de  $y_t$  es  $N(0, \sigma^2 \phi^{-2})$ , pero si  $a_t$  no es Normal, entonces  $y_t$  ya no es independiente.

Se han propuesto otros muchos modelos en la literatura. Granger y Ding (1995) sugieren que las series diarias de rendimientos financieros podrían caracterizarse con la distribución doble exponencial. Ryden et al. (1998) proponen modelizar los rendimientos mediante una mixtura de variables normales de media cero, ya que pueden generar series con muchas de las propiedades empíricas observadas. Ellos proponen usar el modelo de “hidden Markov” para captar la dependencia de mayor orden presente en las series de rendimientos.

Finalmente, los modelos NIGSV (Normal Inverse Gaussian Stochastic Volatility), propuestos por Barndorff-Nielsen (1997) y extendidos y aplicados a series reales por Andersson (2001) son otra alternativa interesante para modelizar series de rendimientos financieros. Recientemente han aparecido muchos artículos relacionados con este tipo de modelos; ver, por ejemplo, Barndorff-Nielsen y Shephard (2001a), Barndorff-Nielsen y Shephard (2001b), Barndorff-Nielsen y Shephard (2002a), Barndorff-Nielsen y Shephard (2002b) y Forsberg y Bollerslev (2002).



2.5 Apéndice de tablas y figuras

Figura 2.1: Rendimientos diarios de varios tipos de cambio

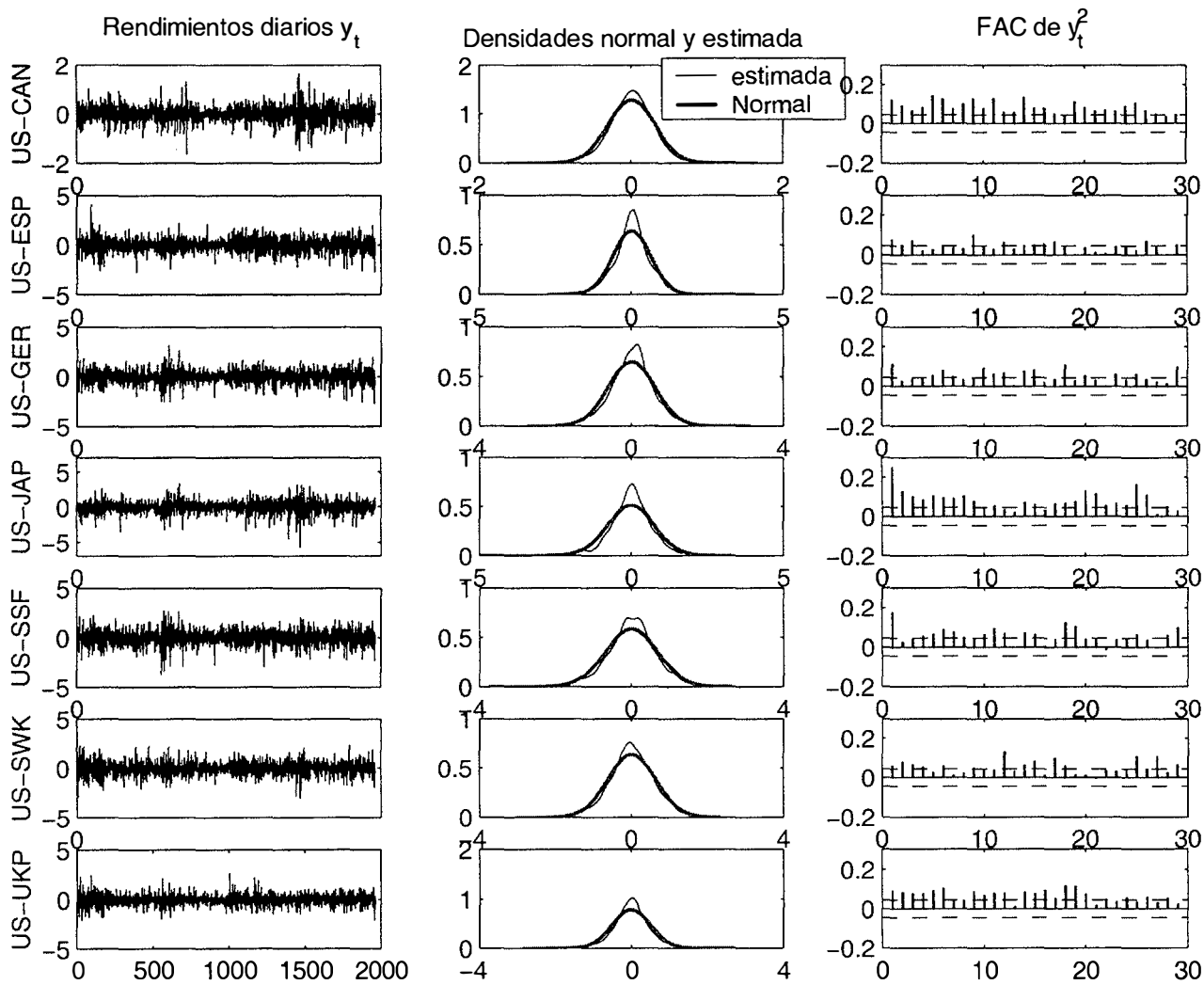


Tabla 2.1: Estadísticos descriptivos de varias series diarias de rendimientos financieros

	US CAN	US ESP	US GER	US JAP	US SSF	US SWK	US UKP	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
T	1963	1963	1963	1963	1963	1963	1963	1261	1262	2728	1982	2888
Media	0.0084	0.0245	0.0153	0.0077	0.0078	0.0149	0.0008	0.0818*	0.0111	0.0472*	0.0708*	0.0575*
D.T.	0.3094	0.6233	0.6129	0.7796	0.6860	0.6267	0.5090	1.3468	1.8471	0.9312	1.2495	0.9085
Simetría	0.1449*	0.1509*	0.2553*	0.5858*	0.3539*	0.1535*	0.0266	0.3101*	0.1329	0.4303*	0.3540*	0.6108*
Curtosis	5.6080*	5.1553*	4.7606*	7.3970*	5.0157*	4.4174*	5.4891*	5.7362*	6.8799*	7.9133*	6.8751*	9.6405*
$r(1)$	0.0647*	0.0341	0.0262	0.0255	0.0278	0.0050	0.0100	0.0300	0.0700*	0.0300	0.1219*	0.0001
$Q(20)$	34.80*	24.20	20.70	23.60	18.70	20.07	32.70*	39.70*	41.30*	47.50*	72.80*	38.30*
Autocorrelaciones de $y_t^2$												
$r_2(1)$	0.1200*	0.0800*	0.1200*	0.2500*	0.1800*	0.0600*	0.0900*	0.1800*	0.1300*	0.2100*	0.2100*	0.1700*
$r_2(2)$	0.0900*	0.0500*	0.0300	0.1300*	0.0300	0.0800*	0.0800*	0.2700*	0.1300*	0.1400*	0.2000*	0.0900*
$r_2(5)$	0.1400*	0.0300	0.0600*	0.1100*	0.0600*	0.0300	0.0900*	0.1300*	0.1000*	0.1600*	0.1800*	0.1500*
$r_2(10)$	0.0800*	0.0400	0.0900*	0.0500	0.0600*	0.0400	0.0700*	0.1600*	0.0400	0.0600*	0.2300*	0.0700*
$Q_2(20)$	380*	96.7*	161*	377*	220*	141*	256*	749*	146*	511*	1414*	497*

T: Tamaño muestral.

$r(\tau)$ : Autocorrelación de orden  $\tau$  de las observaciones  $y_t$ .

$r_2(\tau)$ : Autocorrelación de orden  $\tau$  de los cuadrados  $y_t^2$ .

$Q(20)$  y  $Q_2(20)$ : Estadístico Box Ljung para  $y_t$  e  $y_t^2$  respectivamente (31.4 es el valor crítico al 5%).

\* Significativamente distinto de cero al 95% de confianza.

Figura 2.2: Rendimientos diarios de varios índices bursátiles

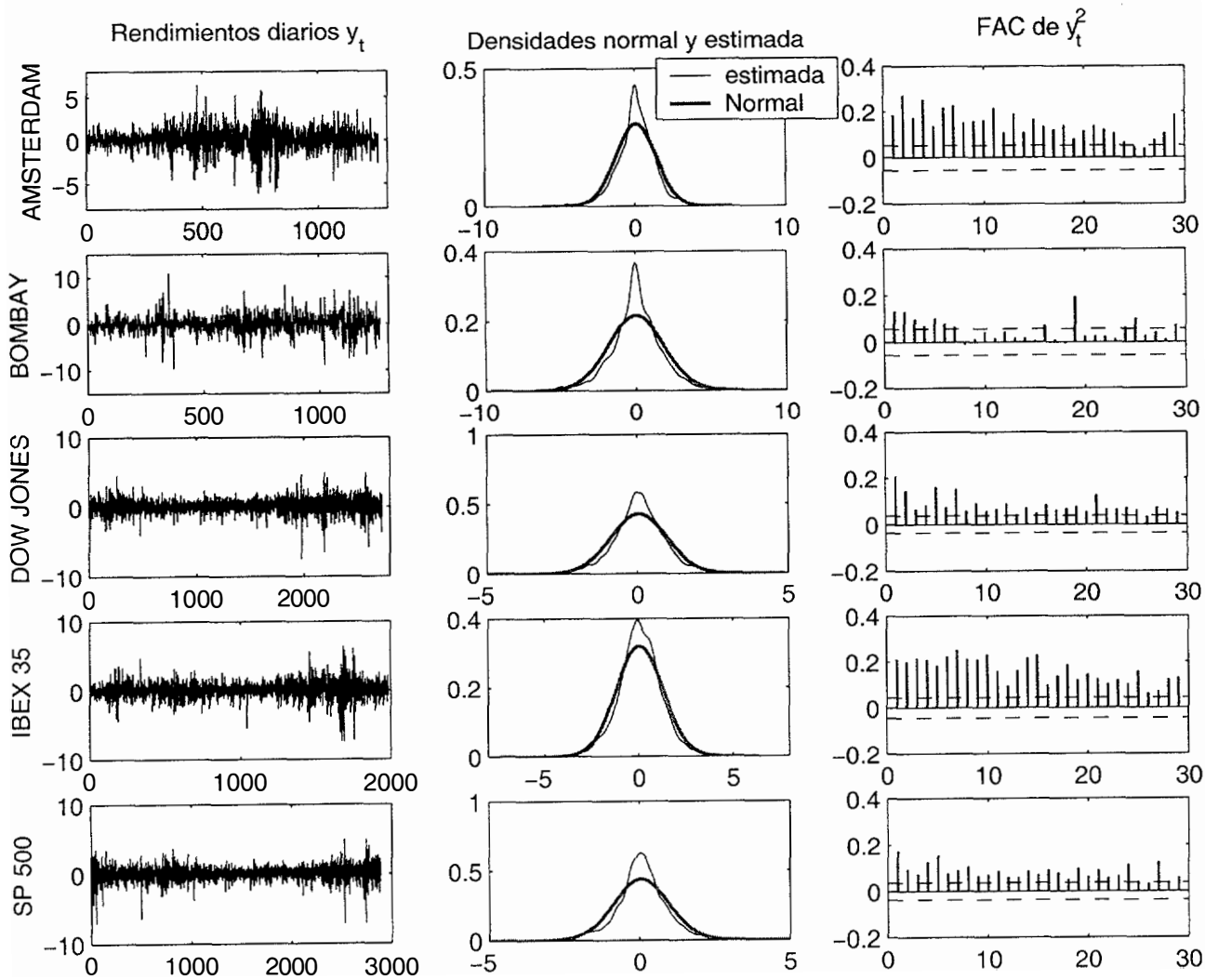


Figura 2.3: Función de Autocorrelación para los rendimientos de S&amp;P 500

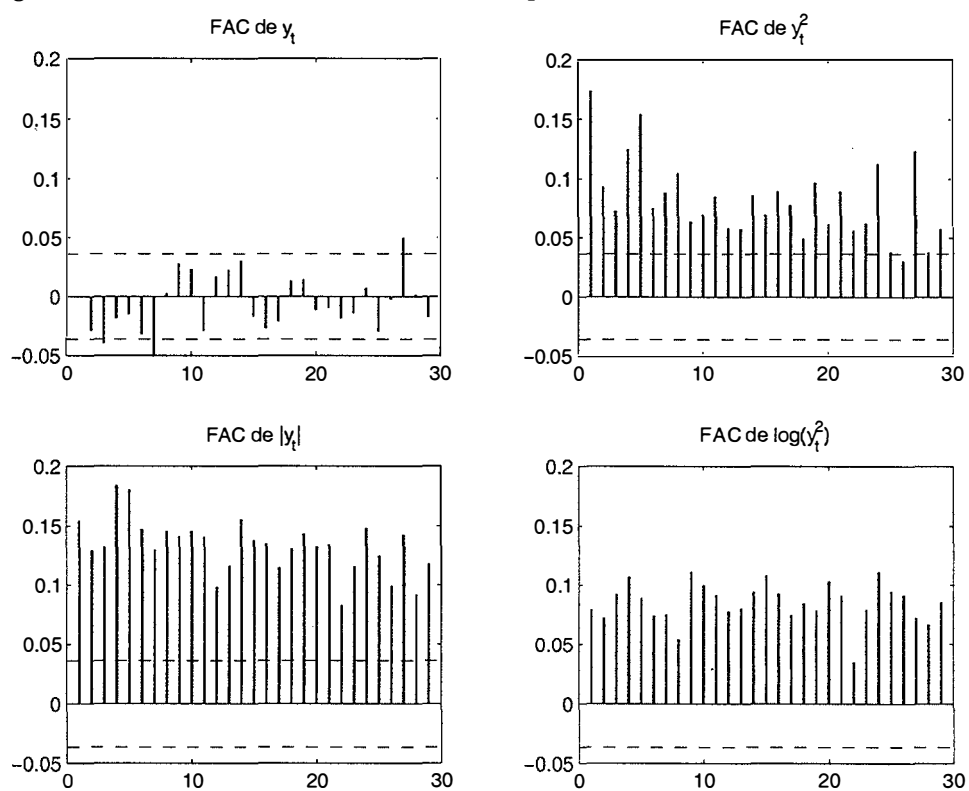


Figura 2.4: Regiones donde el modelo GARCH(1,1) satisface las condiciones de estacionariedad estricta y de existencia de los momentos de órdenes dos y cuatro para distintas distribuciones condicionales

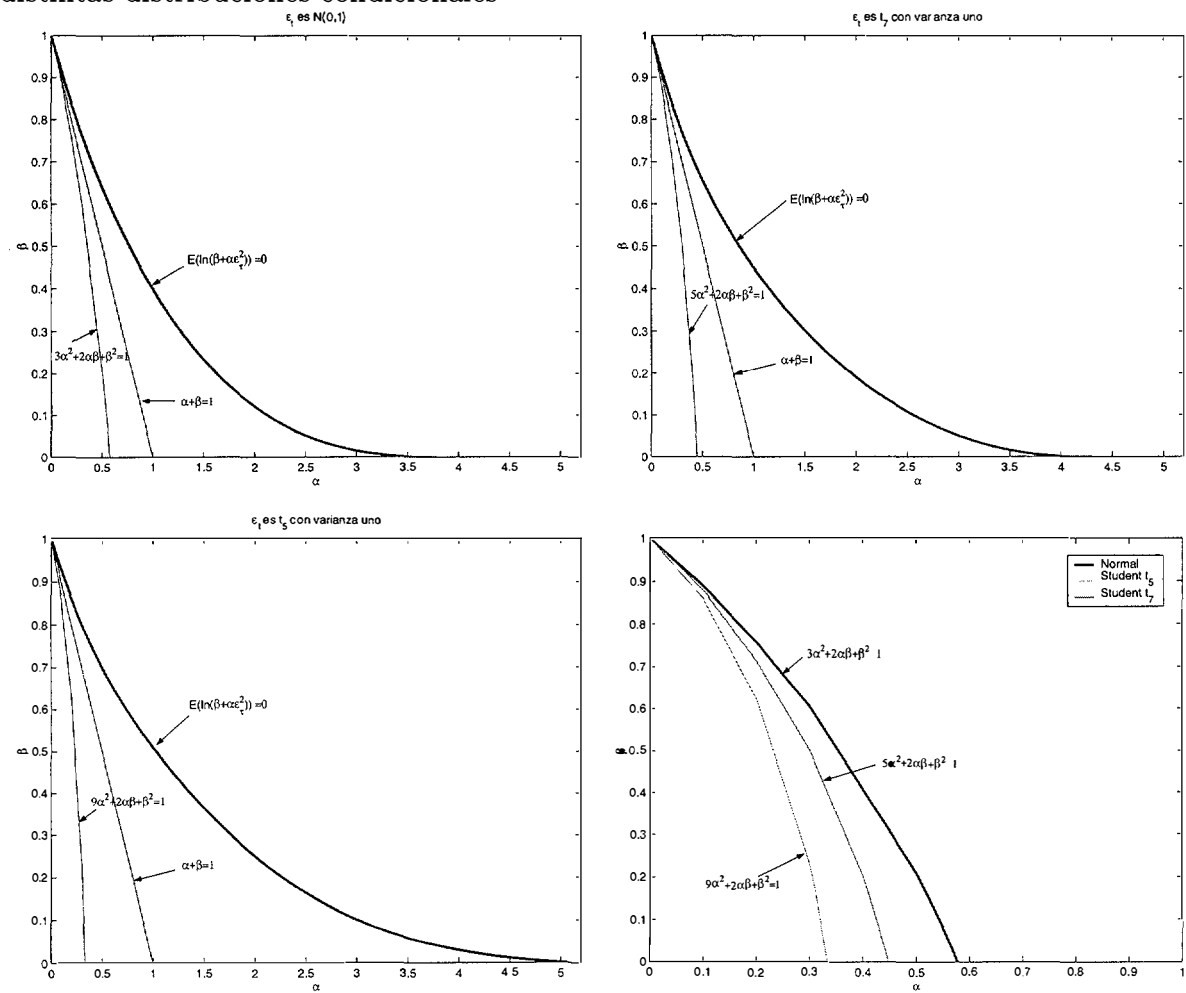


Tabla 2.2: Modelos GARCH(1,1) gaussianos estimados y diagnosis de los residuos

	US CAN	US ESP	US GER	US JAP	US SSF	US SWK	US UKP	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
$\hat{\omega}$	0.0011 (0.001)	0.0045 (0.002)	0.0047 (0.002)	0.0130 (0.002)	0.0116 (0.003)	0.0054 (0.002)	0.0048 (0.001)	0.0174 (0.005)	0.0702 (0.011)	0.0090 (0.002)	0.0470 (0.010)	0.0042 (0.001)
$\hat{\alpha}$	0.0442 (0.005)	0.0328 (0.006)	0.0319 (0.005)	0.0559 (0.006)	0.0301 (0.006)	0.0319 (0.006)	0.0347 (0.005)	0.0659 (0.009)	0.0434 (0.006)	0.0584 (0.005)	0.0979 (0.014)	0.0326 (0.003)
$\hat{\beta}$	0.9449 (0.006)	0.9564 (0.009)	0.9559 (0.008)	0.9230 (0.007)	0.9445 (0.012)	0.9537 (0.010)	0.9443 (0.008)	0.9256 (0.009)	0.9369 (0.006)	0.9323 (0.006)	0.8697 (0.018)	0.9619 (0.003)
$\hat{\alpha} + \hat{\beta}$	0.9891	0.9892	0.9878	0.9789	0.9746	0.9856	0.9790	0.9915	0.9803	0.9907	0.9676	0.9945
log L	379	1808	1772	2162	2000	1814	1364	-2007	2491	3431	3026	-3535
$\hat{\varepsilon}_t = \frac{y_t}{\hat{\sigma}_t}$												
Media	0.0319	0.0381	0.0243	0.0105	0.0110	0.0249	0.0009	0.0113	0.0023	0.0201	0.0198	0.0056
D.T.	1.0013	1.0014	1.0014	1.0000	1.0011	1.0015	0.9987	1.0006	1.0000	1.0006	0.9988	1.0010
Simetría	0.0889	0.1414*	0.3310*	0.5577*	0.4142*	0.1099*	0.0750	0.4051*	0.0637	0.4644*	0.1232*	0.7494*
Curtosis	4.4519*	5.1538*	4.4781*	6.0271*	4.6648*	4.1090*	5.0067*	4.5140*	6.8533*	5.5081*	4.8366*	8.3018*
$r(1)$	0.0050	0.0330	0.0400	0.0270	0.0440	0.0090	0.0280	0.0070	0.0110	0.0090	0.0140	0.0310
$Q(20)$	20.57	18.55	20.68	19.82	19.95	16.94	28.64	21.36	26.50	24.34	19.38	28.63
$r_2(1)$	0.0250	0.0110	0.0360	0.0270	0.0660	0.0250	0.0120	0.0160	0.0500	0.0190	0.0220	0.0300
$r_2(2)$	0.0100	0.0180	0.0210	0.0000	0.0280	0.0070	0.0060	0.0270	0.0170	0.0370	0.0010	0.0190
$r_2(5)$	0.0170	0.0150	0.0120	0.0090	0.0120	0.0240	0.0110	-0.0080	0.0110	0.0070	0.0020	0.0070
$r_2(10)$	0.0150	0.0040	0.0240	0.0270	0.0120	0.0030	0.0230	-0.0030	0.0050	0.0040	0.0030	0.0010
$Q_2(20)$	10.07	7.03	16.87	14.06	20.48	18.23	10.13	11.81	31.49*	11.53	6.98	7.65

Figura 2.5: Estimaciones GARCH(1,1) y ARSV(1) de la volatilidad

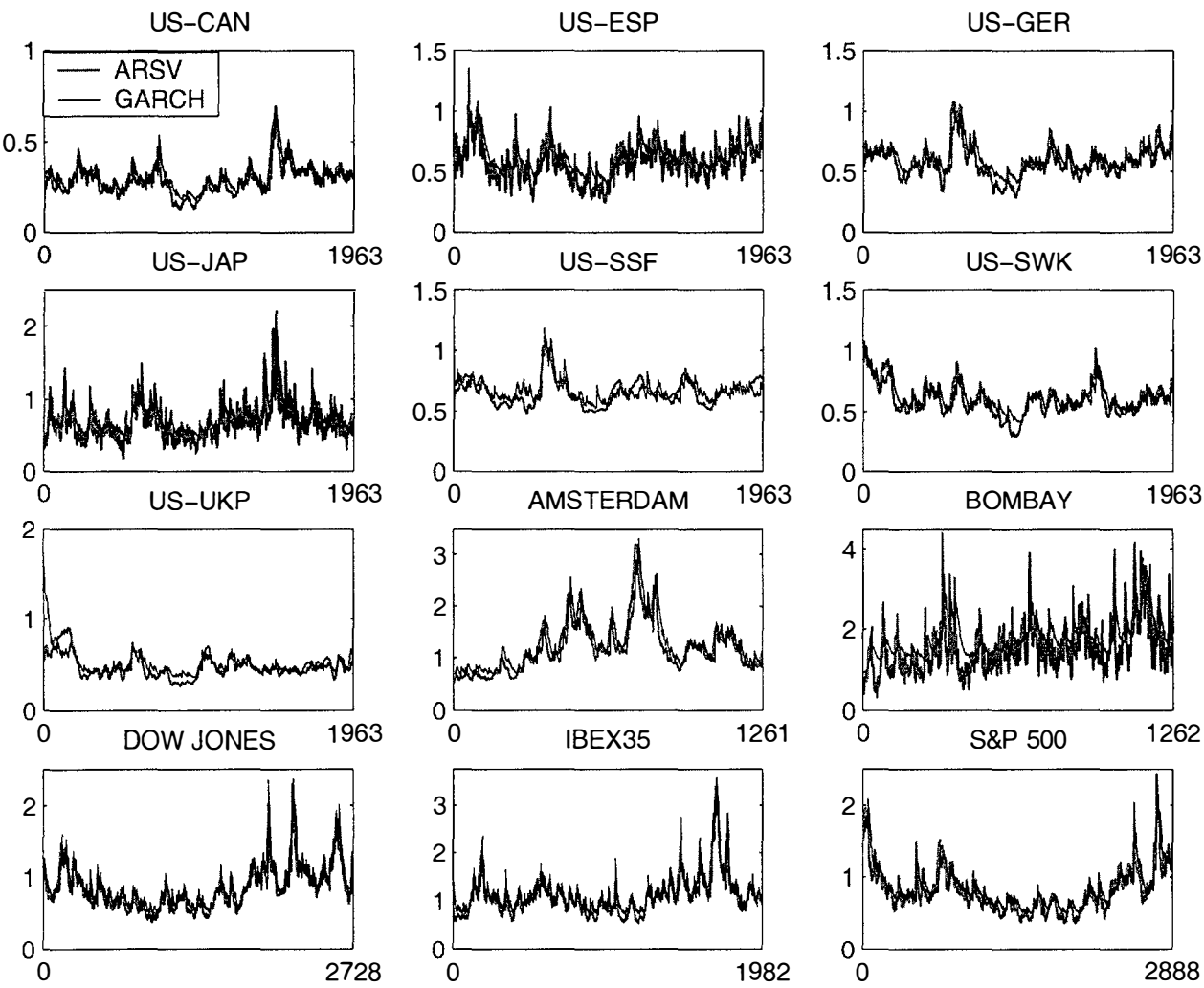


Figura 2.6: Estimaciones kernel de las densidades de las observaciones estandarizadas

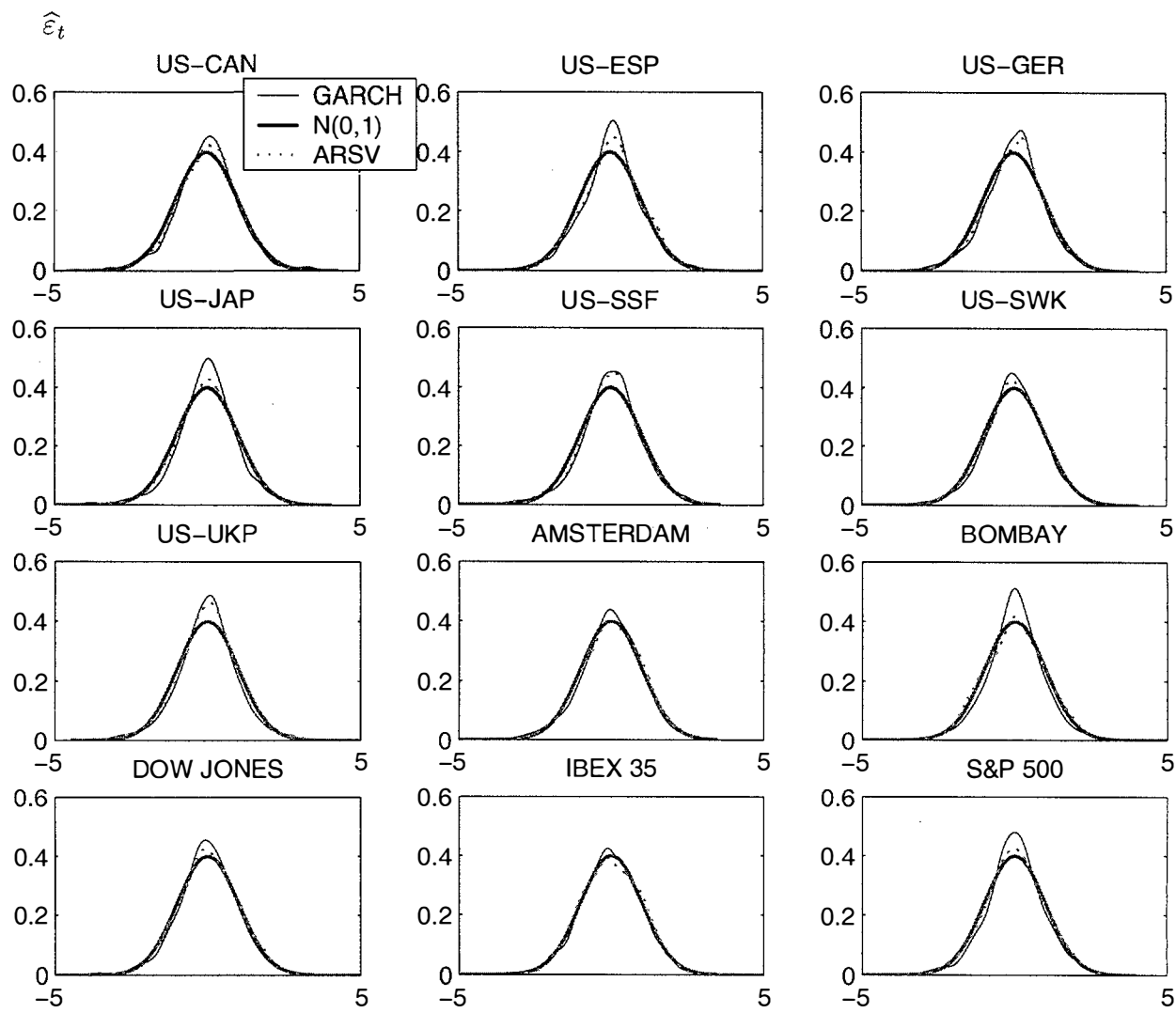




Tabla 2.3: Modelos GARCH(1,1)- $t_\nu$  estimados y diagnosis de los residuos

	US CAN	US ESP	US GER	US JAP	US SSF	US SWK	US UKP	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
$\hat{\omega}$	0.0010 (0.001)	0.0017 (0.001)	0.0025 (0.002)	0.0057 (0.003)	0.0072 (0.004)	0.0020 (0.001)	0.0010 (0.001)	0.0178 (0.008)	0.2663 (0.106)	0.0060 (0.002)	0.0341 (0.011)	0.0026 (0.001)
$\hat{\alpha}$	0.0497 (0.009)	0.0316 (0.006)	0.0364 (0.008)	0.0400 (0.009)	0.0300 (0.008)	0.0418 (0.008)	0.0396 (0.001)	0.0755 (0.015)	0.1317 (0.034)	0.0507 (0.009)	0.0977 (0.017)	0.0360 (0.002)
$\hat{\beta}$	0.9418 (0.011)	0.9655 (0.007)	0.9588 (0.009)	0.9498 (0.012)	0.9558 (0.013)	0.9565 (0.007)	0.9601 (0.001)	0.9121 (0.017)	0.7993 (0.054)	0.9439 (0.010)	0.8789 (0.0212)	0.9637 (0.002)
$\hat{\alpha} + \hat{\beta}$	0.9915	0.9971	0.9952	0.9898	0.9858	0.9983	0.9997	0.9876	0.9386	0.9946	0.9766	0.9997
$\hat{\nu}$	6.1275	5.8635	6.1014	4.7460	5.9279	7.0415	5.5092	7.8301	4.0764	6.4352	10.957	4.4786
log L	169	877	865	1034	978	895	569	994	1206	-1679	1500	-1609
$\hat{\varepsilon}_t = \frac{y_t}{\hat{\sigma}_t}$												
Media	0.0349	0.0380	0.0237	0.0104	0.0106	0.0249	0.0026	0.0029	0.0054	0.0051	0.0018	0.0076
D.T.	0.9918	0.9943	0.9921	1.0111	0.9896	0.9741	0.9632	1.0258	1.0085	0.9991	1.0106	0.9699
Simetría	0.0876	0.1432*	0.3603*	0.5942*	0.4365*	0.0975	0.0822	0.4147*	0.0252	0.4794*	0.1169*	0.7582*
Curtosis	4.5445	5.2624	4.6406	6.2079	4.7585	4.1777	5.3919	4.6452	8.8427	5.5679	4.9941	8.4703
r(1)	0.0820	0.0330	0.0420	0.0270	0.0450	0.0100	0.0340	0.0070	0.0420	0.0290	0.0080	0.0330
Q(20)	30.10	17.38	20.33	20.79	20.36	17.23	28.45	21.41	32.05*	26.94	18.91	29.97
$r_2(1)$	0.0250	0.0090	0.0280	0.0380	0.0660	0.0230	0.0180	-0.0240	0.0120	0.0230	0.0250	0.0290
$r_2(2)$	0.0180	0.0190	0.0230	0.0160	0.0260	-0.0180	0.0100	0.0190	0.0140	0.0400	0.0040	0.0170
$r_2(5)$	0.0100	0.0190	0.0190	0.0020	0.0110	0.0260	0.0080	0.0120	-0.0070	0.0040	0.0060	0.0060
$r_2(10)$	0.0160	0.0040	0.0170	0.0240	0.0100	0.0020	0.0180	0.0300	0.0120	0.0050	0.0020	0.0010
$Q_2(20)$	10.92	6.78	14.46	14.83	19.76	17.37	8.58	9.73	75.95*	12.31	7.49	7.19

Tabla 2.4: Momentos muestrales y los que implican los modelos estimados

$y_t$	Muestral	GARCH(1,1)	GARCH(1,1) t	EGARCH(1,1)	ARSV(1)
US CAN					
Media	0.0084	0.0000	0.0000	0.0000	0.0000
Varianza	0.0957	0.1009	0.1176	0.0009	0.0944
Curtosis	5.6080	3.6595	19.6207	3.6190	4.4480
$\rho_2(1)$	0.1200	0.1223	0.1696	0.1176	0.1352
$\rho_2(2)$	0.0900	0.1210	0.1682	0.1161	0.1307
$\rho_2(5)$	0.1400	0.1171	0.1639	0.1117	0.1180
$\rho_2(10)$	0.0800	0.1108	0.1571	0.1048	0.0999
US ESP					
Media	0.0245	0.0000	0.0000	0.0000	0.0000
Varianza	0.3885	0.4167	0.5862	0.0063	0.3949
Curtosis	5.1553	3.3339	62.1760	3.2653	4.7185
$\rho_2(1)$	0.0800	0.0784	0.1736	0.0705	0.1381
$\rho_2(2)$	0.0500	0.0776	0.1731	0.0693	0.1240
$\rho_2(5)$	0.0300	0.0751	0.1716	0.0661	0.0906
$\rho_2(10)$	0.0400	0.0711	0.1691	0.0611	0.0550
US GER					
Media	0.0153	0.0000	0.0000	0.0000	0.0000
Varianza	0.3757	0.3852	0.5208	0.0099	0.3804
Curtosis	4.7606	3.2748	17.8359	3.2526	4.1474
$\rho_2(1)$	0.1200	0.0704	0.1529	0.0688	0.1157
$\rho_2(2)$	0.0300	0.0695	0.1522	0.0675	0.1102
$\rho_2(5)$	0.0600	0.0670	0.1500	0.0640	0.0953
$\rho_2(10)$	0.0900	0.0630	0.1464	0.0585	0.0752
US JAP					
Media	-0.0077	0.0000	0.0000	0.0000	0.0000
Varianza	0.6078	0.6161	0.5588	0.0277	0.5807
Curtosis	7.3970	3.5281	53.0184	3.5237	5.6888
$\rho_2(1)$	0.2500	0.1202	0.1094	0.1344	0.1701
$\rho_2(2)$	0.1300	0.1176	0.1083	0.1281	0.1518
$\rho_2(5)$	0.1100	0.1103	0.1050	0.1111	0.1093
$\rho_2(10)$	0.0500	0.0992	0.0998	0.0881	0.0654

Tabla 2.5: Momentos muestrales y los que implican los modelos estimados

$y_t$	Muestral	GARCH(1,1)	GARCH(1,1) t	EGARCH(1,1)	ARSV(1)
US-SSF					
Media	0.0078	0.0000	0.0000	0.0000	0.0000
Varianza	0.4706	0.4567	0.5070	0.0619	0.4748
Curtosis	5.0157	3.1124	7.3039	3.1858	3.6090
$\rho_2(1)$	0.1800	0.0469	0.0596	0.0520	0.0761
$\rho_2(2)$	0.0300	0.0457	0.0587	0.0507	0.0744
$\rho_2(5)$	0.0600	0.0423	0.0563	0.0472	0.0695
$\rho_2(10)$	0.0600	0.0372	0.0524	0.0419	0.0621
US-SWK					
Media	0.0149	0.0000	0.0000	0.0000	0.0000
Varianza	0.3928	0.3750	1.1765	0.0056	0.4058
Curtosis	4.4174	3.2299	‡	3.3130	4.0164
$\rho_2(1)$	0.0600	0.0647	0.3667	0.0758	0.1088
$\rho_2(2)$	0.0800	0.0637	0.3660	0.0746	0.1054
$\rho_2(5)$	0.0300	0.0610	0.3642	0.0711	0.0960
$\rho_2(10)$	0.0400	0.0568	0.3611	0.0657	0.0823
US-UKP					
Media	0.0008	0.0000	0.0000	0.0000	0.0000
Varianza	0.2591	0.2286	3.3333	0.0109	0.2717
Curtosis	5.4891	3.1845	‡	3.2055	4.6822
$\rho_2(1)$	0.0900	0.0613	0.7340	0.0610	0.1472
$\rho_2(2)$	0.0800	0.0600	0.7338	0.0598	0.1423
$\rho_2(5)$	0.0900	0.0563	0.7332	0.0563	0.1287
$\rho_2(10)$	0.0700	0.0506	0.7321	0.0510	0.1093
AMSTERDAM					
Media	0.0818	0.0000	0.0000	0.0000	0.0000
Varianza	1.8139	2.0471	1.4355	0.0094	1.6989
Curtosis	5.7362	6.1614	26.0765	3.7926	5.9980
$\rho_2(1)$	0.1800	0.2549	0.2468	0.1475	0.1952
$\rho_2(2)$	0.2700	0.2527	0.2438	0.1439	0.1905
$\rho_2(5)$	0.1300	0.2463	0.2348	0.1336	0.1773
$\rho_2(10)$	0.1600	0.2360	0.2206	0.1183	0.1577

Tabla 2.6: Momentos muestrales y los que implican los modelos estimados

$y_t$	Muestral	GARCH(1,1)	GARCH(1,1)-t	EGARCH(1,1)	ARSV(1)
<b>BOMBAY</b>					
Media	0.0111	0.0000	0.0000	0.0000	0.0000
Varianza	3.4118	3.5635	3.8594	0.6335	3.3492
Curtosis	6.8799	3.3207	#	3.3061	7.4160
$\rho_2(1)$	0.1300	0.0866	0.2238	0.1064	0.1805
$\rho_2(2)$	0.1300	0.0848	0.2083	0.0984	0.1438
$\rho_2(5)$	0.1000	0.0799	0.1681	0.0783	0.0770
$\rho_2(10)$	0.0400	0.0724	0.1176	0.0542	0.0304
<b>DOW JONES</b>					
Media	0.0472	0.0000	0.0000	0.0000	0.0000
Varianza	0.8672	0.9677	1.1111	0.0098	0.8691
Curtosis	7.9133	4.7501	#	4.1482	5.2224
$\rho_2(1)$	0.2100	0.2034	0.2326	0.1709	0.1710
$\rho_2(2)$	0.1400	0.2015	0.2313	0.1660	0.1668
$\rho_2(5)$	0.1600	0.1960	0.2276	0.1525	0.1547
$\rho_2(10)$	0.0600	0.1870	0.2215	0.1327	0.1367
<b>IBEX 35</b>					
Media	0.0708	0.0000	0.0000	0.0000	0.0000
Varianza	1.5613	1.4506	1.4573	0.0105	1.4441
Curtosis	6.8751	4.2899	9.4372	3.7490	5.2052
$\rho_2(1)$	0.2100	0.2116	0.2481	0.1546	0.1692
$\rho_2(2)$	0.2000	0.2047	0.2422	0.1494	0.1638
$\rho_2(5)$	0.1800	0.1854	0.2256	0.1350	0.1487
$\rho_2(10)$	0.2300	0.1573	0.2004	0.1144	0.1271
<b>S&amp;P 500</b>					
Media	0.0575	0.0000	0.0000	0.0000	0.0000
Varianza	0.8253	0.7636	8.6667	0.0016	0.8251
Curtosis	9.6405	3.7210	#	3.8644	5.6730
$\rho_2(1)$	0.1700	0.1176	0.6948	0.1335	0.1851
$\rho_2(2)$	0.0900	0.1169	0.6946	0.1318	0.1797
$\rho_2(5)$	0.1500	0.1150	0.6939	0.1269	0.1645
$\rho_2(10)$	0.0700	0.1119	0.6929	0.1192	0.1426

Figura 2.7: Curva NIC para los modelos GARCH, AVGARCH y EGARCH

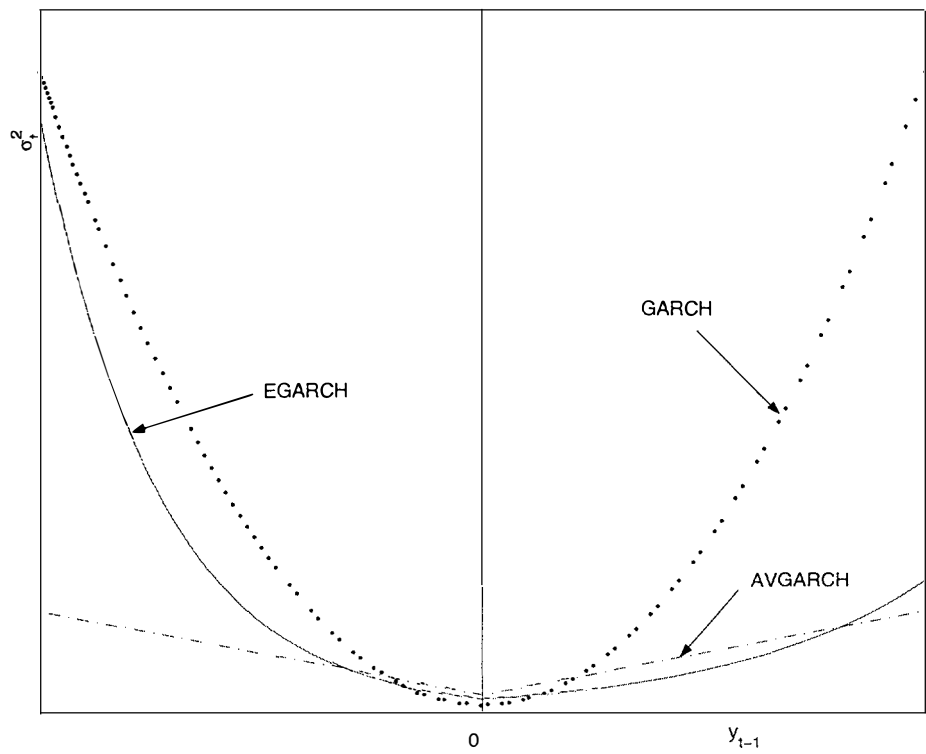


Figura 2.8: FAC de los cuadrados de un modelo EGARCH(1,1) para distintos valores de los parámetros

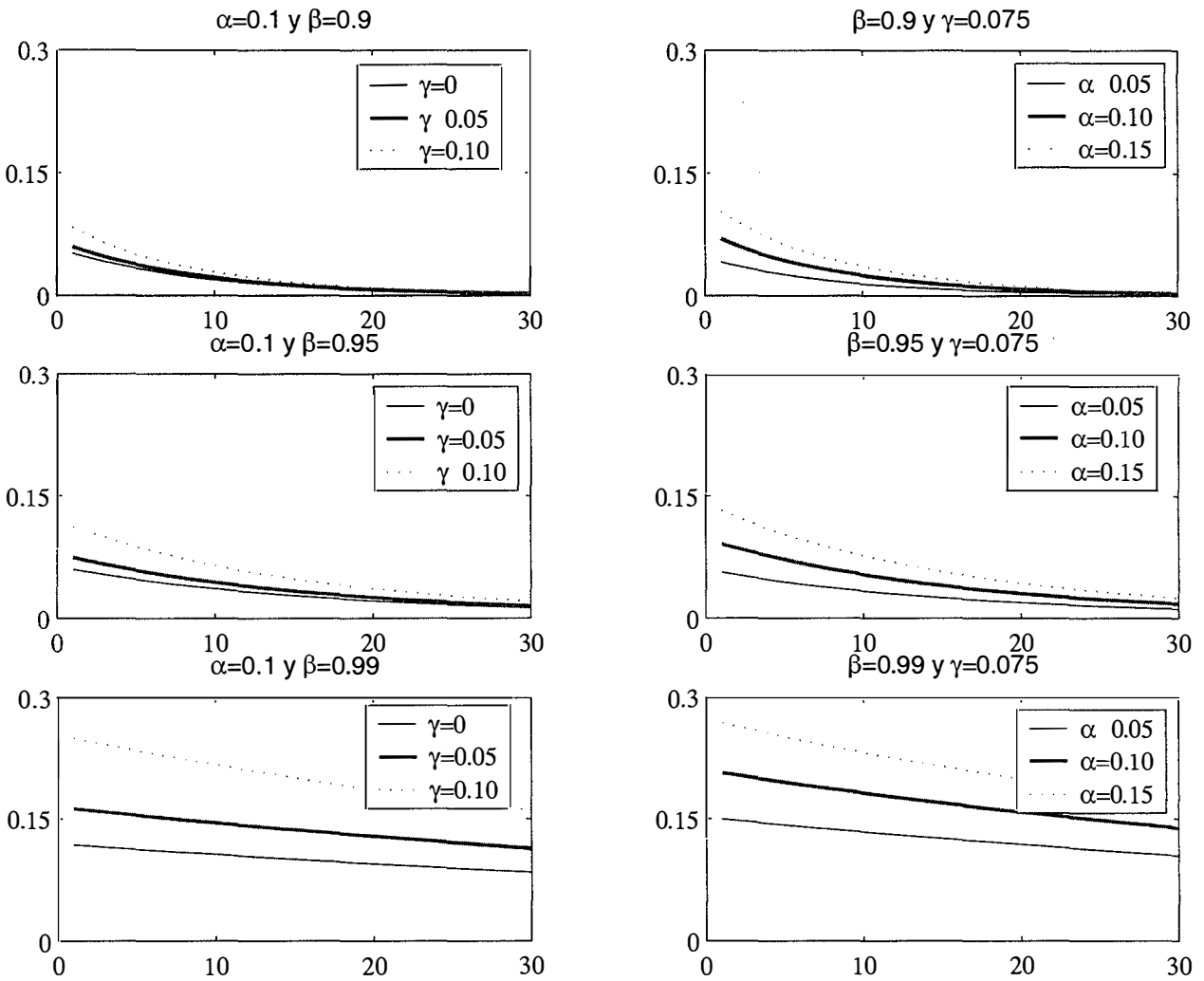


Tabla 2.7: Modelos EGARCH(1,1) gaussianos estimados y diagnosis de los residuos

	US CAN	US ESP	US GER	US JAP	US SSF	US SWK	US UKP	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
$\hat{\omega}$	-0.104 (0.012)	0.074 (0.013)	0.076 (0.013)	0.143 (0.012)	-0.061 (0.013)	-0.074 (0.014)	-0.082 (0.011)	-0.095 (0.013)	-0.032 (0.011)	-0.102 (0.009)	-0.126 (0.015)	-0.067 (0.007)
$\hat{\alpha}$	0.101 (0.011)	0.079 (0.013)	0.079 (0.013)	0.162 (0.013)	0.058 (0.013)	0.077 (0.014)	0.075 (0.011)	0.136 (0.017)	0.150 (0.016)	0.126 (0.011)	0.162 (0.021)	0.086 (0.009)
$\hat{\beta}$	0.989 (0.003)	0.986 (0.005)	0.984 (0.005)	0.961 (0.005)	0.978 (0.006)	0.986 (0.005)	0.982 (0.004)	0.981 (0.004)	0.936 (0.009)	0.979 (0.003)	0.973 (0.006)	0.989 (0.001)
$\hat{\gamma}$	0.014 (0.008)	0.007 (0.006)	-0.015 (0.006)	-0.040 (0.007)	-0.035 (0.006)	0.022 (0.008)	-0.015 (0.006)	-0.041 (0.010)	0.051 (0.011)	-0.079 (0.008)	-0.036 (0.010)	-0.046 (0.006)
log L	376	1802	1769	2156	1996	1807	1367	2005	2489	3400	3025	3509
$\hat{\varepsilon}_t$ $\frac{y_t}{\hat{\sigma}_t}$												
Media	0.0319	0.0380	0.0236	0.0068	0.0099	0.0247	0.0023	0.0035	0.0049	0.0047	0.0167	0.0094
D.T.	1.0012	1.0017	1.0011	1.0003	1.0011	1.0010	0.9996	1.0001	0.9994	1.0003	0.9986	1.0008
Simetría	0.0973	0.1622*	0.3316*	0.4754*	0.4097*	0.0835	0.0705	0.3949*	0.0299	0.3819*	0.1667*	0.6863*
Curtosis	4.4660*	4.9906*	4.5470*	5.7950*	4.6062*	3.9873*	4.8515*	4.7808*	7.6909*	4.9647*	5.1762*	7.8432*
r(1)	0.0000	0.0340	0.0390	0.0310	0.0380	0.0100	0.0280	0.0090	0.0140	0.0060	0.0110	0.0280
Q(20)	22.03	18.82	19.64	20.47	19.50	17.28	30.03	24.03	28.97	22.56	18.66	27.18
r <sub>2</sub> (1)	0.0270	0.0150	0.0320	0.0200	0.0600	0.0280	0.0250	0.0150	0.0230	0.0080	0.0130	0.0210
r <sub>2</sub> (2)	0.0080	0.0180	0.0240	0.0020	0.0300	0.0110	0.0030	0.0300	0.0180	0.0140	0.0090	0.0050
r <sub>2</sub> (5)	0.0250	0.0120	0.0080	0.0210	0.0140	0.0280	0.0260	0.0070	0.0150	0.0100	0.0050	0.0030
r <sub>2</sub> (10)	0.0170	0.0020	0.0290	0.0210	0.0190	0.0030	0.0260	0.0220	0.0020	0.0080	0.0080	0.0050
Q <sub>2</sub> (20)	12.35	7.21	17.96	16.31	18.83	18.95	14.40	12.81	62.43*	8.48	6.62	6.36

Figura 2.9: FAC de los cuadrados de un modelo ARSV(1) para distintos valores de los parámetros

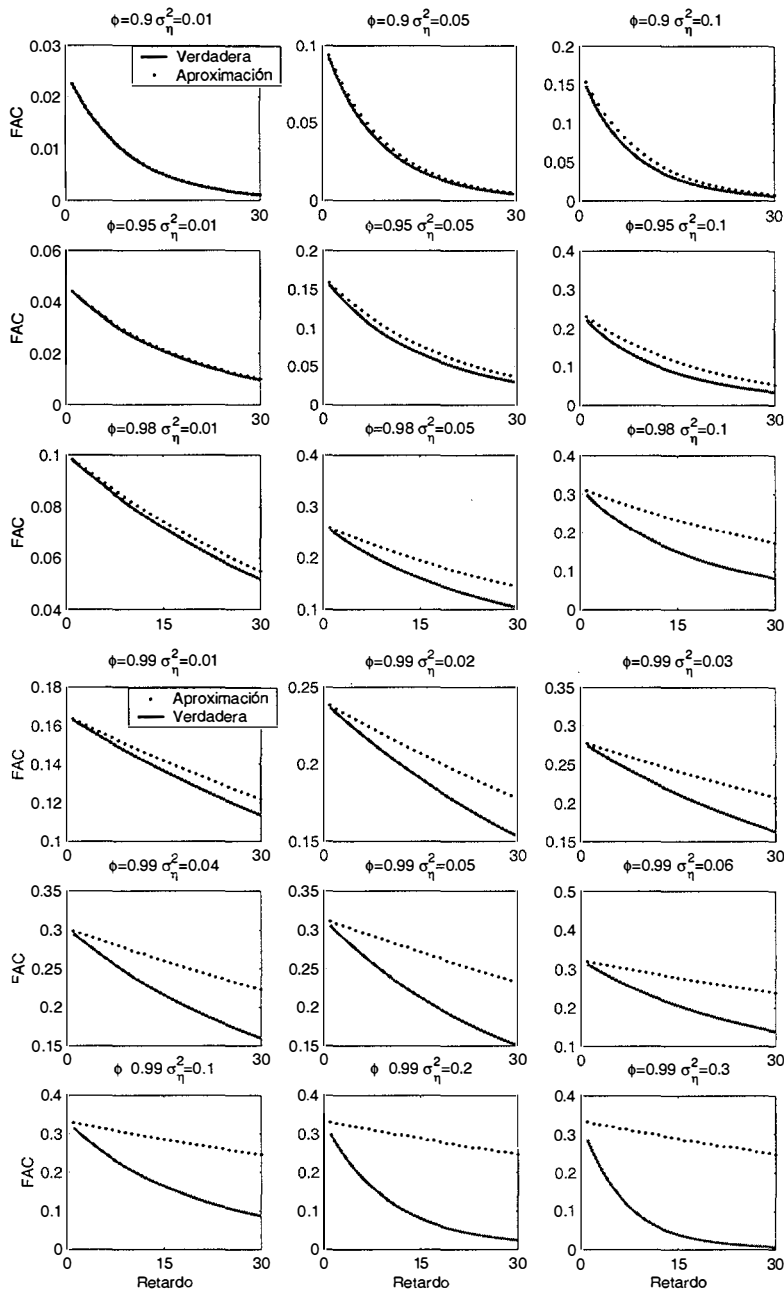




Figura 2.10: Decaimiento de la FAC de los cuadrados de un modelo ARSV para distintos valores de los parámetros

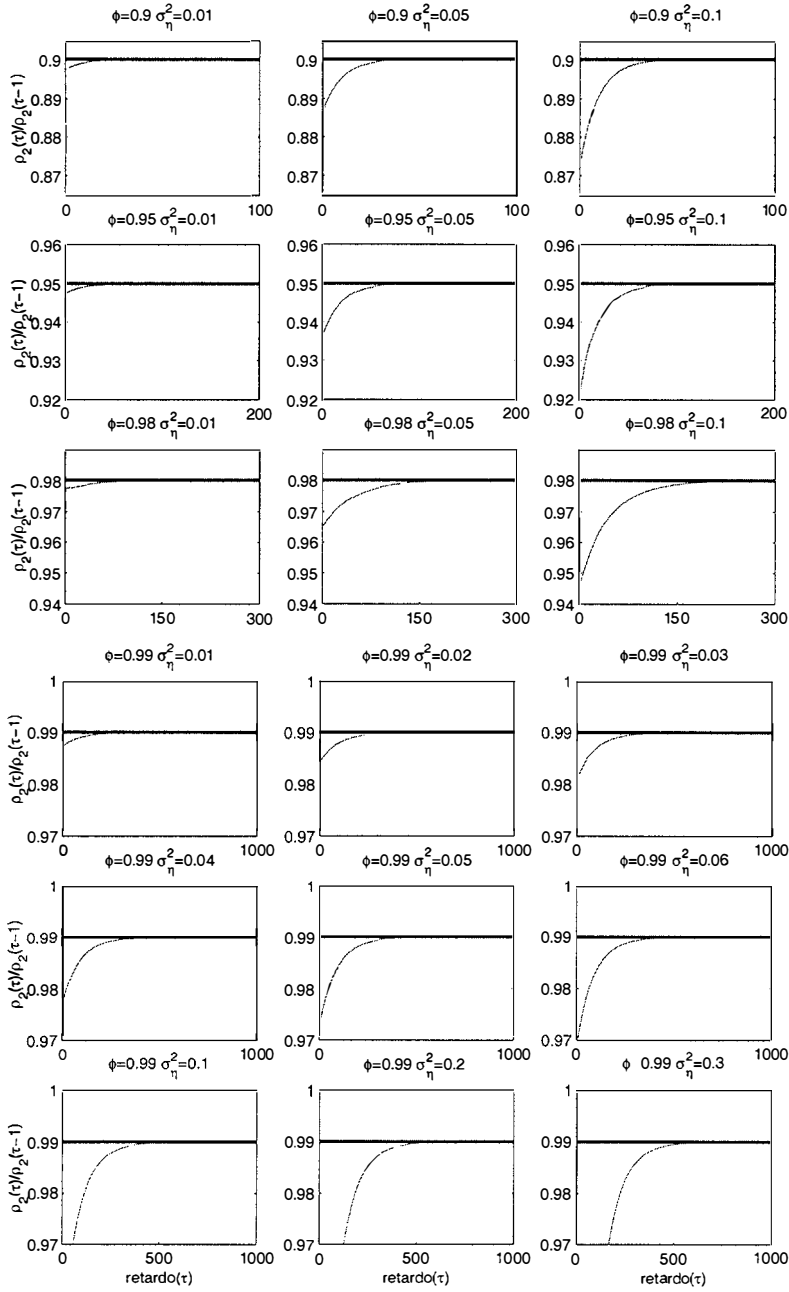


Tabla 2.8: Modelos ARSV(1) gaussianos estimados y diagnosis de los residuos

	US CAN	US ESP	US GER	US JAP	US SSF	US SWK	US UKP	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
$\hat{\phi}$	0.9718 (0.011)	0.9149 (0.038)	0.9587 (0.022)	0.9165 (0.027)	<i>0.9795</i> (0.013)	0.9729 (0.013)	<i>0.9729</i> (0.012)	0.9826 (0.007)	0.8499 (0.041)	0.9806 (0.006)	0.9751 (0.005)	0.9781 (0.007)
$\hat{\sigma}_\eta^2$	0.0219 (0.008)	0.0738 (0.039)	0.0262 (0.016)	0.1024 (0.036)	<i>0.0075</i> (0.006)	0.0156 (0.008)	<i>0.0238</i> (0.012)	0.0239 (0.009)	0.2513 (0.075)	0.0213 (0.006)	0.0271 (0.074)	0.0276 (0.008)
$\hat{\sigma}_*^2$	0.0775 (0.009)	0.3149 (0.026)	0.3235 (0.031)	0.4217 (0.039)	<i>0.4329</i> (0.022)	0.3507 (0.039)	<i>0.2175</i> (0.011)	1.2015 (0.293)	2.1302 (0.226)	0.6587 (0.096)	1.0963 (0.164)	0.6001 (0.087)
Verosim.	0.1806	0.9021	0.8901	1.0692		0.9158		1.5782	1.9142	1.2382	1.5148	1.1945
$\hat{\epsilon}_t$ $\frac{y_t}{\hat{\sigma}_t}$												
Media	0.0342	0.0448	0.0344	0.0143	0.0027	0.0222	0.0054	0.0264	0.0068	0.0152	0.0159	0.0125
D.T.	0.9747	0.9666	0.9741	0.9792	1.0003	0.9782	1.0003	0.9737	0.9710	0.9794	0.9734	0.9743
Simetría	0.0891	0.1214*	0.1676*	0.0951	0.3171*	0.0864	0.0215	0.0929	0.0291	0.0996	0.0035	0.0795
Curtosis	3.2906*	2.9741	3.2417*	3.1853	4.3398*	3.1738	4.4069*	2.9706	2.7658	3.1676	2.7648	3.3363*
$r(1)$	0.0040	0.0430	0.0170	0.0070	0.0323	0.0030	0.0166	0.0010	0.0470	0.0110	0.0100	0.0260
$Q(20)$	25.64	21.11	18	16.21	19.69	16.39	26.61	24.81	30.86	24.43	19.69	30.77
$r_2(1)$	0.0430	0.0540*	0.0400	0.0810*	0.0518	0.0130	0.0247	0.0520	0.0930*	0.0490	0.0620*	0.0520
$r_2(2)$	0.0710*	0.0730*	0.0780*	0.0930*	0.0339	0.0640	0.0228	0.0340	0.0830*	0.0390	0.0250	0.0660*
$r_2(5)$	0.0030	0.0350	0.0390	0.0490	0.0017	0.0590	0.0060	0.0530	0.0080	0.0320	0.0100	0.0100
$r_2(10)$	0.0370	0.0040	0.0180	0.0450	0.0091	0.0240	0.0003	0.0660	0.0190	0.0110	0.0030	0.0130
$Q_2(20)$	48.21*	48.59*	50.35*	57.14*	19.01	39.19*	19.00	26.69	61.53*	41.01*	45.12*	55.33*

Las estimaciones en *cursiva* han sido calculadas con el método de Cuasi Máxima Versosimilitud.

## Capítulo 3

# Persistencia y curtosis en modelos GARCH y ARSV

En este capítulo se analiza la relación entre el coeficiente de curtosis marginal, la persistencia de las innovaciones sobre la volatilidad y la correlación de orden uno de las observaciones al cuadrado en los modelos GARCH(1,1) y ARSV(1). La diferencia entre como se relacionan estos coeficientes en ambos modelos permite explicar dos fenómenos empíricos observados frecuentemente al ajustar dichos modelos a series financieras reales. En particular, se explica por qué, para las mismas series, la persistencia de las innovaciones sobre la volatilidad se estima, a menudo, mayor en modelos GARCH que en modelos ARSV. También se explica por qué los modelos ARSV gaussianos parecen ser adecuados a la hora de explicar las propiedades de las series reales mientras que si el modelo que se ajusta es un GARCH, la distribución de los errores necesita ser leptocúrtica.

### 3.1 Introducción

Las series de rendimientos financieros observadas con frecuencias elevadas se caracterizan principalmente por tener: (1) elevada curtosis marginal, (2) correlación de orden uno de los cuadrados significativa, aunque de magnitud relativamente pequeña y (3) decaimiento muy lento hacia cero, es decir, elevada persistencia en la función de autocorrelación de los rendimientos al cuadrado. Estas tres características empíricas se ponen de manifiesto en la tabla 2.1 y han sido documentadas por distintos autores; ver, por ejemplo, Liesenfeld y Jung (2000) y Loudon et al. (2000).

Los modelos de heterocedasticidad condicional descritos en el capítulo anterior son capaces de generar estos efectos. En este capítulo, nos centramos en los dos modelos paramétricos básicos usados en el análisis empírico de series financieras, el modelo GARCH(1,1) y el modelo ARSV(1).

Se ha observado que la persistencia<sup>1</sup> de las innovaciones sobre la volatilidad, se estima por lo general, mayor en el modelo GARCH(1,1) que en el modelo ARSV(1). Por ejemplo, Taylor (1994) muestra que el parámetro que mide la persistencia en rendimientos diarios del tipo de cambio DM/US Dollar, se estima como 0.9702 en el GARCH(1,1), mientras que es 0.9380 cuando el modelo que se ajusta es el ARSV(1). Resultados similares han sido encontrados por Shephard (1996), Kim et al. (1998), Hafner y Herwartz (2000) y Andersson (2001). Este hecho también se pone de manifiesto en las doce series analizadas en el capítulo anterior. Observando las estimaciones GARCH(1,1) y ARSV(1) en las tablas 2.2 y 2.8, respectivamente, se puede ver que  $\hat{\alpha} + \hat{\beta}$  es mayor que  $\hat{\phi}$  para diez de las doce series. Sin embargo, ninguno de los autores anteriores da una explicación a este fenómeno.

---

<sup>1</sup>El concepto de persistencia en los modelos de heterocedasticidad condicional no está definido con precisión, como ya se ha discutido en el capítulo anterior, sin embargo se va a considerar, como es habitual, que la medida de persistencia en los modelos GARCH(1,1) es  $\alpha + \beta$  y en los modelos ARSV(1) es  $\phi$ .

Dado que la persistencia estimada con los modelos GARCH es, generalmente, muy cercana a uno, en muchas aplicaciones empíricas se asume que la volatilidad tiene una raíz unitaria. Sin embargo, esto implica que la predicción a largo plazo de las volatilidades futuras tiende a infinito de un modo lineal. Puesto que este comportamiento no es muy realista, es importante distinguir cuando la persistencia está cercana a uno y cuando no.

También se ha observado que el supuesto de gaussianidad para las innovaciones,  $\varepsilon_t$ , parece adecuado en los modelos ARSV(1) mientras que, en la especificación GARCH(1,1),  $\varepsilon_t$  requiere una distribución con colas pesadas. Por ejemplo, Shephard (1996) muestra que el modelo GARCH(1,1) con innovaciones que siguen una distribución  $t$  de Student tiene diagnosis similares al modelo ARSV(1) gaussiano. Ghysels et al. (1996), Kim et al. (1998) y Hafner y Herwartz (2000), entre otros, aportan resultados similares. Por lo general, la obtención de las propiedades estadísticas de un modelo y la inferencia son más simples bajo el supuesto de gaussianidad. Por lo tanto, si la distribución gaussiana es una aproximación adecuada, ésta es, en general, preferible a cualquier otra distribución.

En este capítulo se muestra que, aunque los modelos GARCH(1,1) y ARSV(1) son capaces de explicar exceso de curtosis y correlaciones de los rendimientos al cuadrado significativas y con un decaimiento lento, la relación entre la curtosis, la correlación de orden uno de las observaciones al cuadrado y el parámetro asociado con la persistencia que implican estos dos modelos es diferente. Esta diferencia puede explicar los fenómenos descritos anteriormente.

La organización de este capítulo es la siguiente. La sección 2 describe las principales propiedades de los modelos GARCH(1,1) y ARSV(1) y explica por qué sería de esperar que la persistencia que implica el modelo GARCH(1,1) sea mayor que la que implica el ARSV(1) y por qué la distribución condicional de los modelos GARCH necesita colas pesadas. Se analizan también algunas propiedades del modelo EGARCH(1,1)

que sugieren que éste podría ser un modelo intermedio entre el GARCH(1,1) y el ARSV(1) a la hora de representar las tres características descritas anteriormente. Los resultados de esta sección sugieren que el modelo ARSV(1) es más flexible que el modelo GARCH(1,1) en el sentido de que es capaz de representar simultáneamente curtosis elevada, baja correlación de orden uno de las observaciones al cuadrado y alta persistencia de las innovaciones sobre la volatilidad. En la sección 3 se ilustran los resultados mediante una serie de experimentos de Monte Carlo. Finalmente, las conclusiones aparecen en la sección 4.

## 3.2 Propiedades de los modelos GARCH y ARSV

### 3.2.1 El modelo GARCH(1,1)

Teräsvirta (1996) observó que los valores de la curtosis y de la correlación de orden uno que implica el modelo GARCH(1,1) con errores gaussianos no se corresponden con los observados en series reales. En particular, valores altos de la curtosis están asociados a valores de la correlación de orden uno de los cuadrados, mayores que los usualmente observados en la práctica. Ante este hecho, Teräsvirta (1996) sugirió que distribuciones con colas más pesadas que la distribución normal para  $\varepsilon_t$  podrían mejorar la capacidad de los modelos GARCH(1,1) para caracterizar los hechos observados en la práctica. Para ello estudió la relación que implica el modelo GARCH(1,1) entre la curtosis,  $\kappa_y$  en (2.5) y  $\rho_2(1)$  en (2.7). Sin embargo, dicha relación depende también de la persistencia,  $p = \alpha + \beta$ , por lo que son tres los coeficientes a tener en cuenta. Es posible ver, después de algunos cálculos, que la relación entre  $\kappa_y$ ,  $\rho_2(1)$  y  $p$  viene dada por la expresión:

$$\rho_2(1) = \frac{\sqrt{\frac{(\kappa_y - \kappa_\varepsilon)(1-p^2)}{(\kappa_\varepsilon - 1)\kappa_y}} \left[ 1 - p^2 + p \sqrt{\frac{(\kappa_y - \kappa_\varepsilon)(1-p^2)}{(\kappa_\varepsilon - 1)\kappa_y}} \right]}{1 - p^2 + \frac{(\kappa_y - \kappa_\varepsilon)(1-p^2)}{(\kappa_\varepsilon - 1)\kappa_y}} \quad (3.1)$$

La figura 3.1 representa esta relación para tres distribuciones alternativas de  $\varepsilon_t$ , las distribuciones gaussiana y t de Student con 7 y 5 grados de libertad. Nótese que en los tres casos, dado un valor de la curtosis, la correlación de orden uno disminuye a medida que se incrementa la persistencia. Por lo tanto, el modelo GARCH(1,1) gaussiano sólo es capaz de generar, al mismo tiempo, valores altos de la curtosis y valores bajos de  $\rho_2(1)$  si la persistencia es próxima a uno. Esto sugiere que al ajustar modelos GARCH(1,1) gaussianos,  $\alpha + \beta$  podría estimarse muy próximo a uno, incluso si la persistencia no fuese elevada.

La figura 3.2 representa la relación entre  $\kappa_y$  y  $\rho_2(1)$  para varios modelos GARCH(1,1) con diferente persistencia, junto con los valores muestrales de la curtosis y la correlación de orden uno de los rendimientos al cuadrado para las doce series analizadas en el Capítulo 2 y descritas en la tabla 2.1. Como se puede ver en este gráfico, el modelo GARCH(1,1) con errores gaussianos no es capaz de representar los valores observados. Sin embargo, los modelos GARCH(1,1) con una distribución condicional t de Student, son mejores a la hora de explicar simultáneamente los tres hechos empíricos observados en series temporales reales. Como ilustración, consideremos  $\alpha + \beta = 0.95$  y  $\kappa_y = 10$ . En este caso, si  $\varepsilon_t$  fuese gaussiano o una t de Student con 7 o 5 grados de libertad,  $\rho_2(1)$  sería 0.3831, 0.2037 o 0.0462 respectivamente. Para  $\alpha + \beta = 0.99$  y  $\kappa_y = 10$ , si  $\varepsilon_t$  fuese gaussiano o una t de Student con 7 o 5 grados de libertad,  $\rho_2(1)$  sería 0.3185, 0.1543 o 0.0278 respectivamente. Sin embargo, una distribución t de Student para los errores,  $\varepsilon_t$ , impone restricciones importantes para garantizar la existencia del momento de cuarto orden o hacer inferencia. En particular, cuando los grados de libertad son menores que 8, la distribución asintótica de las correlaciones de los cuadrados ya no es conocida; véase, por ejemplo, McLeod y Li (1983).

Por otra parte, la curtosis de  $y_t$  está muy directamente relacionada con la curtosis de  $\varepsilon_t$ ; ver Bai et al. (2000) para una expresión de esta relación. Por ejemplo, si  $\varepsilon_t$  se distribuye como una t de Student con 7 grados de libertad, la curtosis debe ser mayor

que 5 para que la fac de los cuadrados sea positiva.

### 3.2.2 El modelo EGARCH(1,1)

Debido a la complejidad de las expresiones (2.16) y (2.17) para el coeficiente de curtosis marginal y la fac de las observaciones al cuadrado en el modelo EGARCH(1,1) respectivamente, es muy difícil obtener, como en el modelo GARCH(1,1), una expresión que relacione la curtosis, la correlación de orden uno de los cuadrados y la persistencia en estos modelos. No obstante, es posible representar, para un valor fijo de los parámetros, la fac de las observaciones al cuadrado que implica el modelo EGARCH(1,1). La figura 3.3 representa, para el modelo EGARCH(1,1) simétrico, es decir,  $\gamma = 0$ , dicha fac para un valor determinado de la curtosis y de la persistencia junto con las correspondientes fac que implican los modelos GARCH(1,1) y ARSV(1). Como se puede observar en este gráfico, la fac que implica el modelo EGARCH es diferente de las otras dos. A pesar de que los valores de los coeficientes de correlación no distan mucho de los correspondientes al modelo GARCH(1,1), cabe señalar que el coeficiente de correlación de orden uno es el mayor de los tres modelos y que la tasa a la que decae la fac no es geométrica, como en el GARCH(1,1), sino que es más similar al decaimiento de la fac en el modelo ARSV(1).

Se ha calculado también la relación que implica el modelo EGARCH(1,1) entre la curtosis y el coeficiente de correlación de orden uno de los cuadrados para distintos valores de los parámetros. Las figuras 3.4 y 3.5 representan dicha relación para modelos EGARCH simétricos y asimétricos. Como se puede observar en ambos gráficos, la relación entre  $\kappa_y$  y  $\rho_2(1)$  no es muy diferente de la que implica el modelo GARCH(1,1) aunque parece que el parámetro de asimetría  $\gamma$  permite al modelo generar series con un campo más amplio de valores para la curtosis y la correlación de orden uno de los cuadrados. De nuevo, parece que los valores de  $\rho_2(1)$  asociados a los valores de la



curtosis comúnmente observados en la práctica son mayores que los muestrales y que a medida que el parámetro  $\beta$ , asociado a la persistencia, se acerca a uno, para un valor dado de  $\rho_2(1)$ , los valores de la curtosis son mayores y, para un valor dado de  $\kappa_y$ , los valores de  $\rho_2(1)$  son menores, al igual que pasa en el modelo GARCH.

### 3.2.3 El modelo ARSV(1)

Basándose en la aproximación propuesta por Taylor (1986), dada por (2.23) y discutida en detalle en el capítulo anterior, se ha argumentado frecuentemente que la estructura de las correlaciones de los cuadrados en los modelos GARCH(1,1) y ARSV(1) es muy similar. Por ello, y dado que  $\alpha + \beta$  y  $\phi$  pueden interpretarse como medidas de persistencia en los modelos GARCH(1,1) y ARSV(1) respectivamente, muchos autores han comparado sus estimaciones; ver, por ejemplo, Taylor (1994) y Shephard (1996). No obstante, tanto en la tabla 3.1 como en la figura 3.3, se observa que las fac de las observaciones al cuadrado en modelos GARCH(1,1) y ARSV(1) gaussianos son bastante diferentes. La tabla 3.1 muestra las fac respectivas para distintos valores del coeficiente de curtosis y para  $\phi = \alpha + \beta$ . Como puede observarse, para  $\phi = \alpha + \beta$  y para la misma curtosis, las autocorrelaciones de las observaciones al cuadrado que implica el modelo ARSV(1) son sistemáticamente menores que las que implica el modelo GARCH(1,1)<sup>2</sup>, excepto cuando la volatilidad se acerca a la región de no estacionariedad, donde ambas fac son aproximadamente iguales. Es decir, los dos modelos tienen las mismas propiedades estadísticas cuando la volatilidad está cerca de tener una raíz unitaria. Sin embargo, nótese que en este caso, los parámetros que generan la heterocedasticidad condicional,  $\alpha$  y  $\sigma_\eta^2$  en los modelos GARCH(1,1) y ARSV(1) respectivamente,

---

<sup>2</sup>En base a la figura 3.3, parece que las correlaciones que implica el modelo ARSV(1) en las observaciones al cuadrado son también sistemáticamente menores que las que implica el modelo EGARCH(1,1). No obstante, es necesario estudiar más detalladamente el modelo EGARCH(1,1) para entender mejor cual es la relación que este modelo implica entre las tres características de interés.

están muy cerca de cero. Consecuentemente, en el caso límite de la raíz unitaria, en la práctica, las series temporales generadas por ambos modelos son aproximadamente homocedásticas. Finalmente, nótese que, en los modelos ARSV(1), para un valor dado de la curtosis, la correlación de primer orden de los cuadrados aumenta a medida que el parámetro que mide la persistencia,  $\phi$ , crece. Esta relación es contraria a la observada en modelos GARCH(1,1).

Siguiendo la idea de Teräsvirta (1996), Liesenfeld y Jung (2000) derivan la relación entre curtosis y  $\rho_2(1)$  en un modelo ARSV(1), que viene dada por

$$\rho_2(1) = \frac{\left(\frac{\kappa_y}{\kappa_e}\right)^\phi - 1}{\kappa_y - 1} \quad (3.2)$$

La figura 3.6 representa dicha relación, teniendo en cuenta además la persistencia, junto con la correspondiente relación para el modelo GARCH(1,1), ambos con errores normales. La caja representa los valores usualmente observados en series reales, que son curtosis entre 5 y 10 y  $\rho_2(1)$  entre 0.1 y 0.2. Este gráfico muestra que el modelo ARSV(1) es capaz de generar series que tengan una elevada curtosis, un valor bajo de  $\rho_2(1)$  y una persistencia lejos de la región de no estacionariedad. Sin embargo, las series generadas por modelos GARCH(1,1), han de tener una persistencia muy alta ya que es el único modo de lograr al mismo tiempo valores altos de la curtosis y bajos de  $\rho_2(1)$ . Consecuentemente, el modelo ARSV(1) no necesita distribuciones condicionales con colas anchas o volatilidades con persistencias cercanas a la raíz unitaria para generar valores elevados de la curtosis y valores pequeños de  $\rho_2(1)$ . La presencia de la innovación  $\eta_t$  permite que los modelos ARSV(1) sean más flexibles en el sentido de que pueden generar series con propiedades similares a las observadas en la realidad para un rango más amplio de los parámetros. Este hecho podría explicar por qué, en la práctica, cuando la volatilidad se modeliza con un proceso GARCH(1,1), es necesario suponer una distribución leptocúrtica para  $\varepsilon_t$ , mientras que si se especifica un modelo ARSV(1), el supuesto de gaussianidad para  $\varepsilon_t$  parece adecuado. La figura 3.6 explica también por

qué la persistencia estimada en modelos ARSV(1) es, por lo general, menor que la estimada en modelos GARCH(1,1).

Manteniendo la persistencia,  $\phi$ , constante, es posible representar la relación en (3.2) para los modelos ARSV(1) y GARCH(1,1) junto con los valores de la curtosis y la correlación de orden uno de los rendimientos al cuadrado para las doce series analizadas. La figura 3.7 representa esta relación. Como se puede ver en este gráfico, el modelo ARSV(1) gaussiano es más adecuado que el GARCH(1,1) gaussiano a la hora de representar los valores observados.

Para ilustrar que los modelos GARCH(1,1) con errores  $t$  de Student tienen ajustes similares a los modelos ARSV(1) con errores gaussianos, la figura 3.8 representa la relación entre  $\kappa_y$ ,  $\rho_2(1)$  y persistencia para los modelos ARSV(1) con errores normales y los modelos GARCH(1,1) con errores  $t$  de Student con 7 grados de libertad. En este caso ambos modelos son capaces de generar series con valores de la curtosis y  $\rho_2(1)$  dentro de la caja, donde se encuentran los valores típicamente encontrados en la práctica. No obstante, es importante señalar que mientras que los modelos ARSV(1) pueden generar series con curtosis menores que 5, los modelos GARCH(1,1)- $t_7$  no.

Manteniendo la persistencia constante se obtiene la figura 3.9, que representa la relación entre  $\kappa_y$  y  $\rho_2(1)$  para los modelos ARSV(1) con errores normales y los modelos GARCH(1,1) con errores  $t$  de Student con 7 grados de libertad. Como se puede observar, ambos modelos son capaces de explicar los valores observados.

Finalmente, como en el caso de los modelos GARCH(1,1), los modelos ARSV(1) pueden generar todavía mayores curtosis si los errores,  $\varepsilon_t$ , tienen una distribución leptocúrtica; ver, por ejemplo, Harvey et al. (1994), Chib et al. (1998) y Liesenfeld y Jung (2000). Las figuras 3.10 y 3.11 representan la relación entre  $\kappa_y$ ,  $\rho_2(1)$  y  $\phi$ , y  $\kappa_y$  y  $\rho_2(1)$  respectivamente, para modelos ARSV(1) gaussianos y con errores  $t$  de Student. Se puede observar que suponiendo una distribución  $t$  de Student con un número rel-



ativamente pequeño de grados de libertad, por ejemplo 5, el modelo ARSV(1) genera series con valores muy grandes de la curtosis y muy bajos de  $\rho_2(1)$ , menores que los que se observan usualmente, sólo cuando la persistencia aumenta, el modelo ARSV(1)-t aumenta el valor de  $\rho_2(1)$ . Esto podría explicar por qué los estimadores de  $\phi$  bajo la especificación ARSV(1)-t son, por lo general, mayores que los mismos bajo la especificación ARSV(1)-normal; ver Mahieu y Schotman (1998) y Liesenfeld y Jung (2000).

### 3.3 Experimentos de Monte Carlo

Con objeto de ilustrar el hecho de que los modelos GARCH(1,1) pueden estimar mayor persistencia de la real y que necesitan distribuciones condicionales no gaussianas incluso cuando  $\varepsilon_t$  es gaussiano, se han simulado 1000 series de tamaño muestral  $T=5000$ , generadas por un modelo ARSV(1) gaussiano, a las que se les ha ajustado un modelo GARCH(1,1) gaussiano<sup>3</sup>. La tabla 3.2 presenta los modelos ARSV(1) usados para generar las series y los resultados del experimento de Monte Carlo al estimar los modelos GARCH(1,1). En particular, se muestra la mediana y la desviación típica muestral obtenidas de las 1000 réplicas de los estimadores<sup>4</sup>. Se muestra también el porcentaje de series en las que los parámetros estimados no satisfacen la condición para la existencia de la curtosis,  $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$ . Para esas series, la mediana de la curtosis que implica el modelo ha sido calculada basada en los estimadores que satisfacen dicha condición. La última columna muestra la mediana de  $\rho_2(1)$  que implica el modelo.

Como se puede observar en la tabla 3.2, la persistencia que implica el modelo GARCH(1,1) estimado es mayor que la persistencia real de las series, especialmente para las series con mayor curtosis. Hafner y Herwartz (2000) obtuvieron un resulta-

<sup>3</sup>Los parámetros han sido estimados mediante Cuasi-Máxima Verosimilitud con Fortran 90. Todos los programas están disponibles.

<sup>4</sup>Aunque los resultados son similares para la media, la razón para presentar la mediana es la asimetría de las estimaciones kernel de la densidad de los estimadores.

do similar en un experimento de Monte Carlo con un diseño más restringido, llevado a cabo con un objetivo diferente. Sin embargo, el resultado más interesante que se extrae de la tabla 3.2 es que, para un elevado número de series, los parámetros estimados correspondientes al modelo GARCH(1,1) no satisfacen la condición de existencia de la curtosis. Por ejemplo, en el caso en que  $\phi = 0.9856$  y  $\sigma_\eta^2 = 0.0344$  (véase la fila ocho de dicha tabla), que es bastante realista, la curtosis teórica del proceso está perfectamente definida y es igual a 7. Sin embargo, cuando se estima un modelo GARCH(1,1) gaussiano a estas series, ninguno de los modelos estimados satisface la condición de existencia de la curtosis. Otras dos filas de la tabla (las filas 11 y 12) se corresponden con series con curtosis finita para las que, en ningún caso, los modelos estimados satisfacen la condición de existencia del momento de cuarto orden. Este es un resultado importante que explica por qué la utilización masiva de los procesos GARCH para modelizar series de rendimientos financieros ha llevado a plantear que en los rendimientos financieros la curtosis no está definida. Los resultados en la tabla 3.2 sugieren que el hecho de que dicha condición no se satisfaga puede ser un problema del propio modelo GARCH, que no es capaz de representar adecuadamente las propiedades de series cuya curtosis está perfectamente definida. Finalmente, nótese que aún en los casos en los que la curtosis del modelo estimado existe, los valores estimados y reales de la misma difieren substancialmente. Nótese también que el coeficiente  $\rho_2(1)$  que implica el modelo GARCH(1,1) estimado es siempre mayor que el valor teórico.

Como ilustración, la tabla 3.3 muestra los resultados de la estimación de modelos GARCH(1,1) ajustados a cuatro series concretas generadas por modelos ARSV(1) gaussianos. Esta tabla muestra además los diagnósticos de las observaciones estandarizadas  $\hat{\varepsilon}_t = y_t / \hat{\sigma}_t$ . Como se puede observar, la persistencia estimada,  $\hat{\alpha} + \hat{\beta}$ , es muy próxima a uno y mayor que el valor teórico,  $\phi$ , en los cuatro casos. Por ejemplo, en la cuarta serie, el valor estimado de  $\alpha + \beta$  es 0.9959 cuando la serie había sido generada con  $\phi = 0.98$ . Nótese también que, para las series que satisfacen la condición de existencia

del momento de cuarto orden, las curtosis que implican los modelos estimados son 3.8305 y 5.1604, menores que los valores teóricos, que eran 5.0781 y 6.0000 respectivamente. Para las otras dos series cuya curtosis teórica es 7 y 10.6041, los estimadores GARCH(1,1) correspondientes no satisfacen la condición para la existencia de la curtosis. En lo que a  $\rho_2(1)$  se refiere, el coeficiente que implica el modelo siempre es mayor que el verdadero valor. Por otra parte, aunque todas las series han sido generadas con  $\varepsilon_t$  gaussiano, las observaciones estandarizadas  $\hat{\varepsilon}_t$  no parecen ser gaussianas, puesto que tienen exceso de curtosis. Los diagnósticos sobre las observaciones estandarizadas de la tabla 3.3, son similares a los encontrados en el análisis empírico de series temporales financieras ajustadas por modelos GARCH(1,1); ver, por ejemplo, Andersson (2001) y Kim et al. (1998).

### 3.4 Conclusiones

En este capítulo, se ha discutido la relación entre curtosis, autocorrelación de orden uno de las observaciones al cuadrado y persistencia en modelos GARCH(1,1) y ARSV(1). Aunque los modelos GARCH y ARSV son capaces de explicar exceso de curtosis y autocorrelaciones de los cuadrados significativas y con un decaimiento lento, la relación existente entre la curtosis, el coeficiente de orden uno de la fac de las observaciones al cuadrado y la persistencia que implica cada modelo es diferente. Mientras que los modelos GARCH(1,1) necesitan valores de la persistencia,  $\alpha + \beta$ , muy próximos a uno para que la correlación de orden uno de las observaciones al cuadrado sea pequeña y, al mismo tiempo, la curtosis sea elevada, los modelos ARSV(1) no necesitan imponer valores de la persistencia,  $\phi$ , próximos a uno para generar series con valores de la curtosis grandes y, al mismo tiempo, valores pequeños de la correlación de orden uno de las observaciones al cuadrado. Esto implica que es posible que la persistencia tan elevada que, frecuentemente, se estima con los modelos GARCH(1,1) sea espuria.

---

Alternativamente, en los modelos GARCH es necesario asumir distribuciones condicionales leptocúrticas para poder generar series con curtosis mayores sin incrementar las autocorrelaciones de las observaciones al cuadrado.

En cuanto al modelo EGARCH(1,1), es más flexible que el GARCH(1,1) gracias al parámetro de asimetría, siendo un modelo intermedio entre los modelos GARCH(1,1) y ARSV(1).

3.5 Apéndice de tablas y figuras

Figura 3.1: Relación que implica el modelo GARCH(1,1) entre la curtosis, la correlación de primer orden de los rendimientos al cuadrado y la persistencia

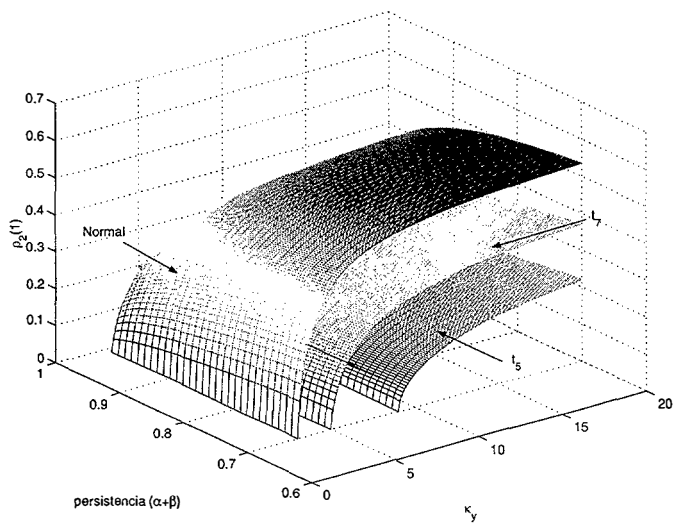


Figura 3.2: Relación que implica el modelo GARCH(1,1) entre la curtosis y la correlación de primer orden de los rendimientos al cuadrado

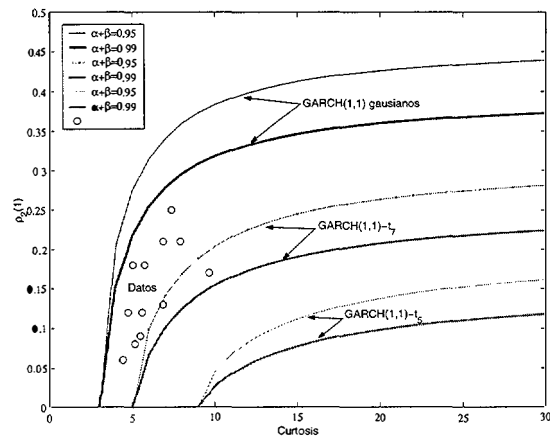




Figura 3.3: Función de autocorrelación de las observaciones al cuadrado que implican los modelos GARCH(1,1), EGARCH(1,1) simétricos y ARSV(1) para un valor dado de la curtosis y de la persistencia

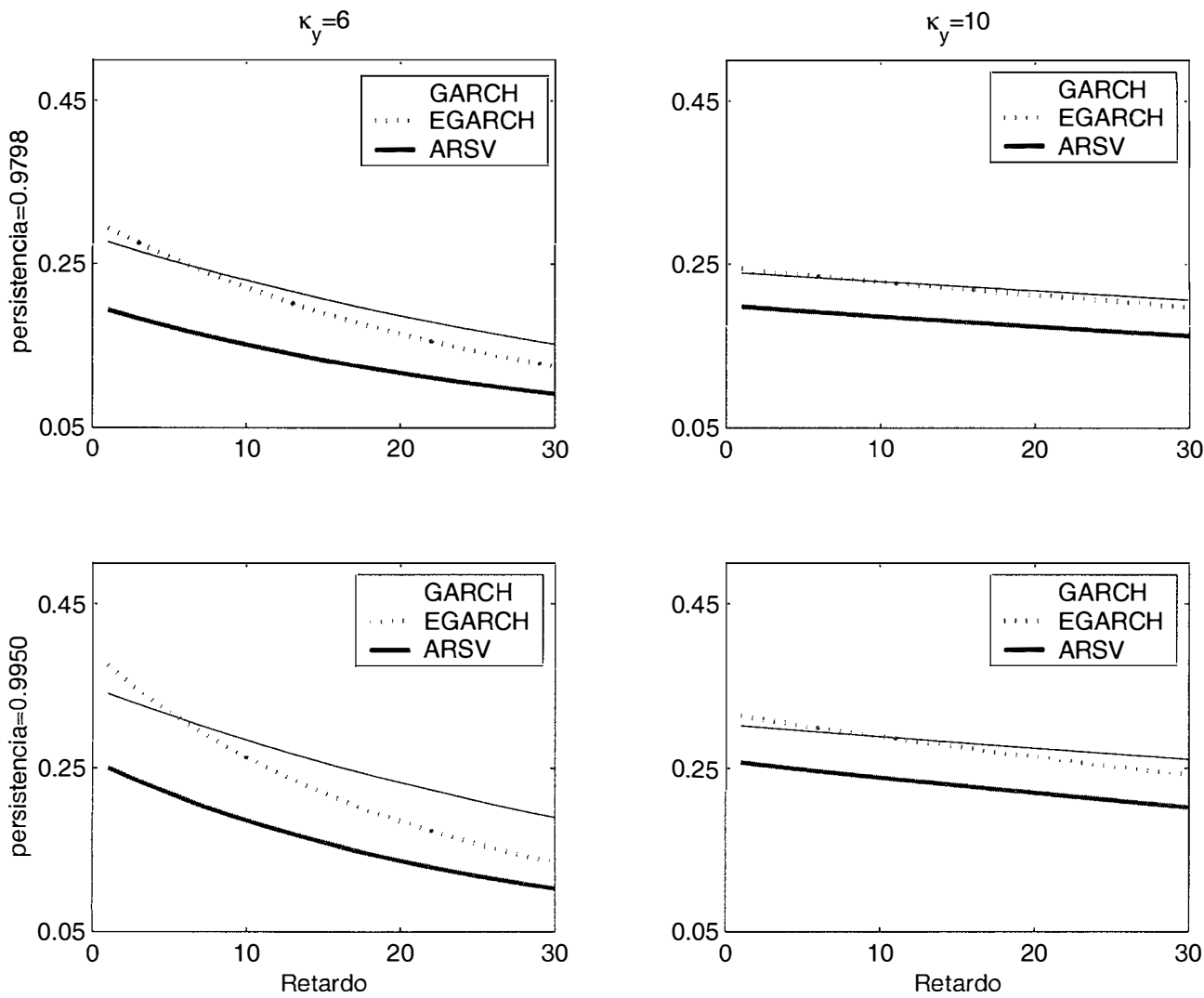


Figura 3.4: Relación que implican los modelos GARCH(1,1) y EGARCH(1,1) simétricos ( $\gamma = 0$ ) entre la curtosis y la correlación de primer orden de los rendimientos al cuadrado

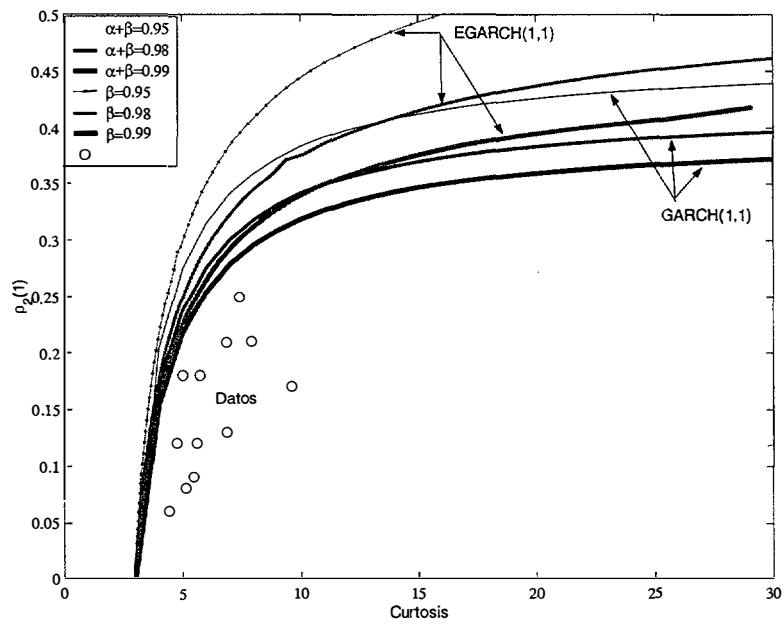


Figura 3.5: Relación que implican los modelos GARCH(1,1) y EGARCH(1,1) asimétricos ( $\gamma \neq 0$ ) entre la curtosis y la correlación de primer orden de los rendimientos al cuadrado

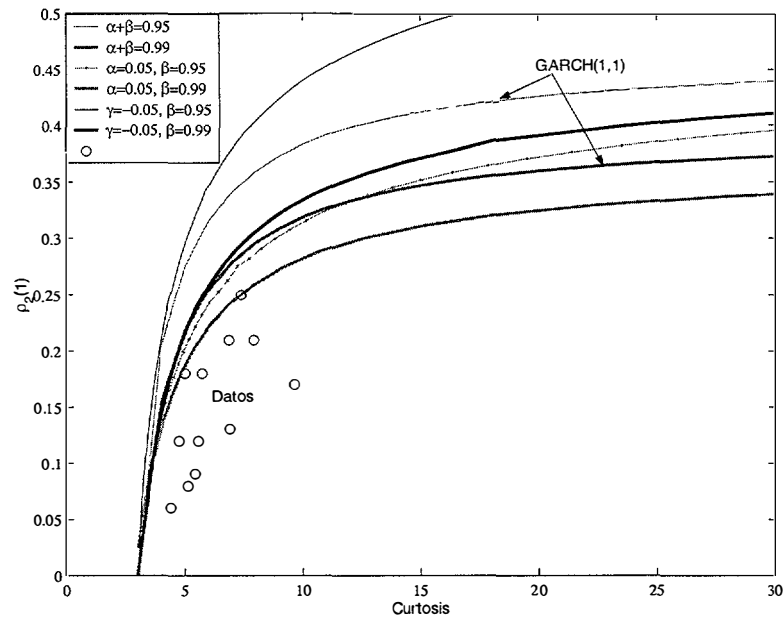


Tabla 3.1: Función de autocorrelación de las observaciones al cuadrado en modelos GARCH(1,1) y ARSV(1) gaussianos

	$\kappa_y = 4$		$\kappa_y = 5$		$\kappa_y = 6$		$\kappa_y = 10$	
	GARCH	ARSV	GARCH	ARSV	GARCH	ARSV	GARCH	ARSV
	$\alpha = 0.2000$ $\beta = 0.6246$	$\sigma_\eta^2 = 0.0919$ $\phi = 0.8250$	$\alpha = 0.2000$ $\beta = 0.6944$	$\sigma_\eta^2 = 0.1016$ $\phi = 0.8950$	$\alpha = 0.2000$ $\beta = 0.7165$	$\sigma_\eta^2 = 0.1109$ $\phi = 0.9165$	$\alpha = 0.2000$ $\beta = 0.7411$	$\sigma_\eta^2 = 0.1401$ $\phi = 0.9400$
$\rho_2(1)$	0.2694	0.0893	0.3157	0.1449	0.3433	0.1775	0.3921	0.2334
$\rho_2(2)$	0.2222	0.0721	0.2824	0.1264	0.3146	0.1580	0.3691	0.2108
$\rho_2(3)$	0.1832	0.0584	0.2526	0.1106	0.2884	0.1410	0.3473	0.1909
$\rho_2(4)$	0.1511	0.0475	0.2259	0.0970	0.2643	0.1262	0.3269	0.1733
$\rho_2(5)$	0.1246	0.0387	0.2021	0.0852	0.2422	0.1131	0.3076	0.1577
	$\alpha = 0.1000$ $\beta = 0.8592$	$\sigma_\eta^2 = 0.0231$ $\phi = 0.9590$	$\alpha = 0.1000$ $\beta = 0.8747$	$\sigma_\eta^2 = 0.0252$ $\phi = 0.9750$	$\alpha = 0.1000$ $\beta = 0.8798$	$\sigma_\eta^2 = 0.0277$ $\phi = 0.9798$	$\alpha = 0.1000$ $\beta = 0.8856$	$\sigma_\eta^2 = 0.0344$ $\phi = 0.9856$
$\rho_2(1)$	0.1955	0.1059	0.2458	0.1614	0.2760	0.1944	0.3296	0.2529
$\rho_2(2)$	0.1875	0.1010	0.2396	0.1563	0.2704	0.1891	0.3249	0.2467
$\rho_2(3)$	0.1798	0.0963	0.2335	0.1514	0.2649	0.1839	0.3202	0.2408
$\rho_2(4)$	0.1725	0.0918	0.2276	0.1467	0.2596	0.1788	0.3156	0.2350
$\rho_2(5)$	0.1654	0.0876	0.2218	0.1421	0.2543	0.1740	0.3110	0.2293
	$\alpha = 0.0500$ $\beta = 0.9399$	$\sigma_\eta^2 = 0.0057$ $\phi = 0.9900$	$\alpha = 0.0500$ $\beta = 0.9437$	$\sigma_\eta^2 = 0.0102$ $\phi = 0.9900$	$\alpha = 0.0500$ $\beta = 0.9450$	$\sigma_\eta^2 = 0.0069$ $\phi = 0.9950$	$\alpha = 0.0500$ $\beta = 0.9464$	$\sigma_\eta^2 = 0.0086$ $\phi = 0.9964$
$\rho_2(1)$	0.1544	0.1098	0.2073	0.1645	0.2390	0.1986	0.2954	0.2577
$\rho_2(2)$	0.1529	0.1086	0.2060	0.1624	0.2378	0.1972	0.2943	0.2561
$\rho_2(3)$	0.1514	0.1073	0.2047	0.1604	0.2366	0.1959	0.2933	0.2545
$\rho_2(4)$	0.1498	0.1061	0.2034	0.1584	0.2354	0.1945	0.2922	0.2529
$\rho_2(5)$	0.1483	0.1049	0.2021	0.1564	0.2342	0.1932	0.2912	0.2514
	$\alpha = 0.0050$ $\beta = 0.9949$	$\sigma_\eta^2 = 5.75 \times 10^{-5}$ $\phi = 0.9999$	$\alpha = 0.0050$ $\beta = 0.9949$	$\sigma_\eta^2 = 10^{-4}$ $\phi = 0.9999$	$\alpha = 0.0050$ $\beta = 0.9949$	$\sigma_\eta^2 = 1.38 \times 10^{-4}$ $\phi = 0.9999$	$\alpha = 0.0050$ $\beta = 0.9949$	$\sigma_\eta^2 = 2.4 \times 10^{-4}$ $\phi = 0.9999$
$\rho_2(1)$	0.1155	0.1111	0.1708	0.1666	0.2040	0.2000	0.2630	0.2592
$\rho_2(2)$	0.1155	0.1111	0.1708	0.1666	0.2040	0.1999	0.2629	0.2592
$\rho_2(3)$	0.1155	0.1111	0.1708	0.1666	0.2040	0.1999	0.2629	0.2591
$\rho_2(4)$	0.1151	0.1111	0.1708	0.1666	0.2040	0.1999	0.2629	0.2591
$\rho_2(5)$	0.1150	0.1110	0.1708	0.1666	0.2039	0.1999	0.2629	0.2590

Figura 3.6: Relación que implican los modelos ARSV(1) y GARCH(1,1) gaussianos entre la curtosis, la correlación de primer orden de los rendimientos al cuadrado y la persistencia

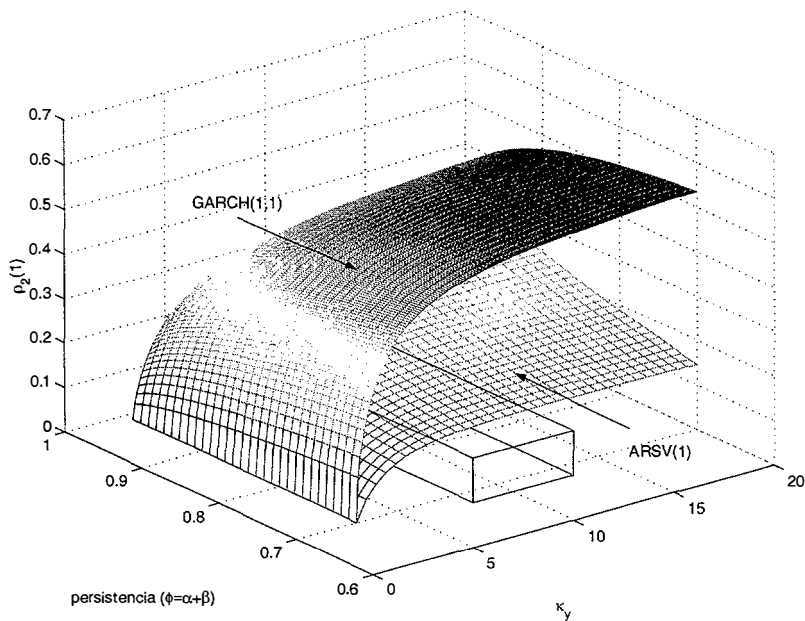


Figura 3.7: Relación que implican los modelos ARSV(1) y GARCH(1,1) gaussianos entre la curtosis y la correlación de primer orden de los rendimientos al cuadrado

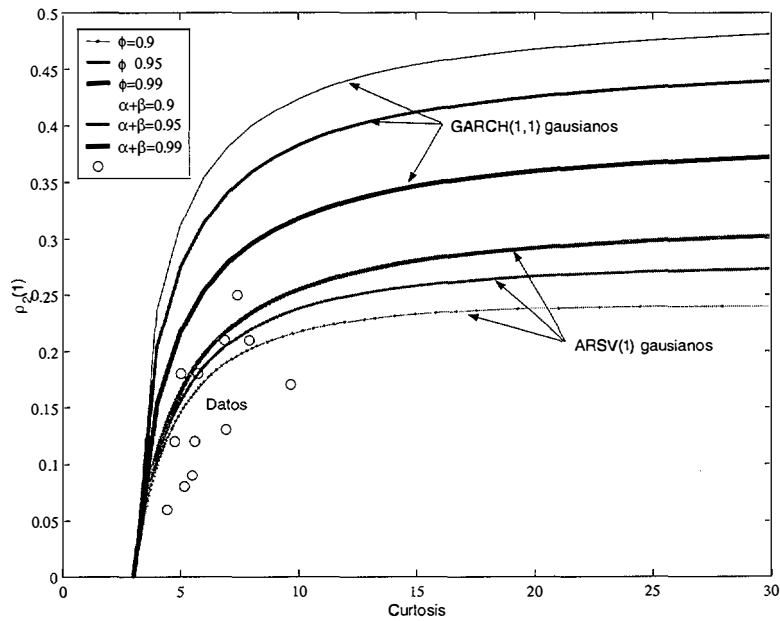


Figura 3.8: Relación que implican los modelos ARSV(1) gaussiano y GARCH(1,1)- $t_7$  entre la curtosis, la correlación de primer orden de los rendimientos al cuadrado y la persistencia

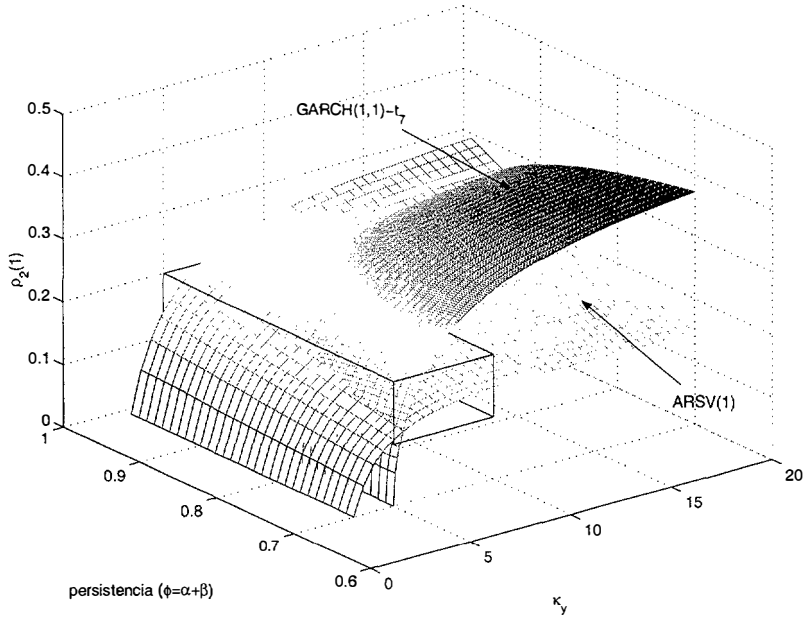


Figura 3.9: Relación que implican los modelos ARSV(1) gaussiano y GARCH(1,1)- $t_7$  entre la curtosis y la correlación de primer orden de los rendimientos al cuadrado

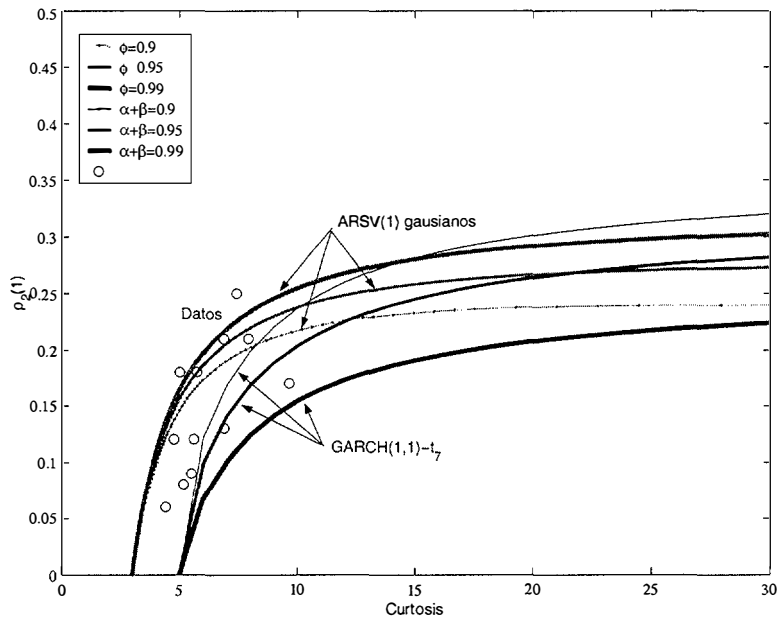


Figura 3.10: Relación que implica el modelo ARSV(1) entre la curtosis, la correlación de primer orden de los rendimientos al cuadrado y la persistencia

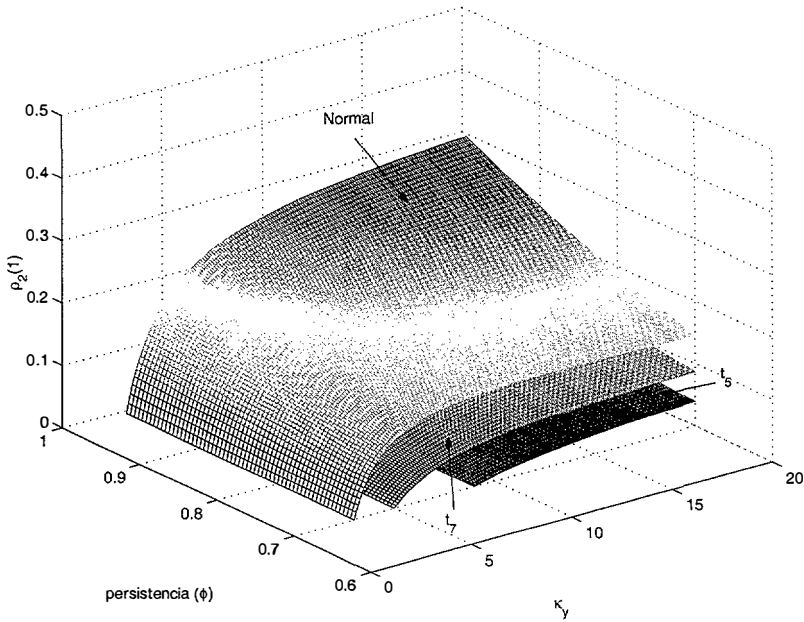


Figura 3.11: Relación que implica el modelo ARSV(1) entre la curtosis y la correlación de primer orden de los rendimientos al cuadrado

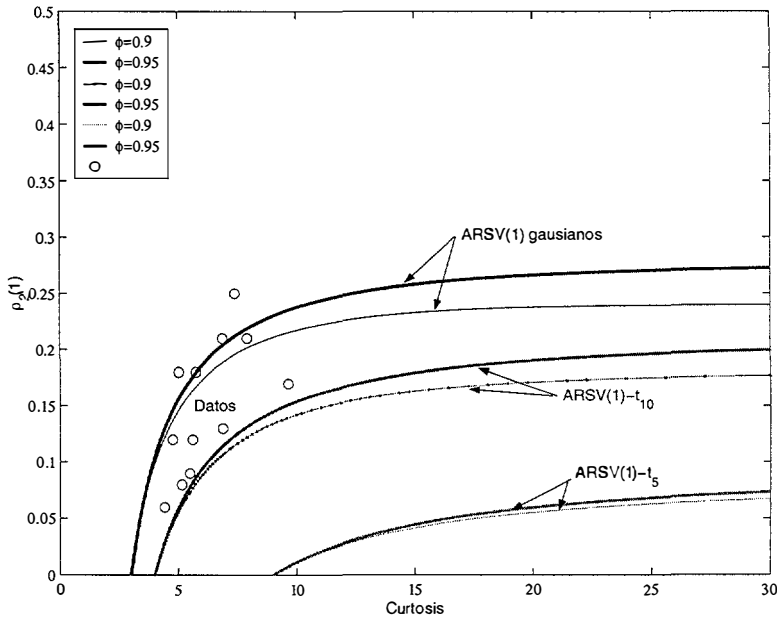


Tabla 3.2: Mediana de los modelos GARCH(1,1) gaussianos estimados a series generadas por modelos ARSV(1) gaussianos basados en 1000 réplicas

MODELO SIMULADO				MODELO ESTIMADO						
ARSV(1)				GARCH(1,1)						
$\phi$	$\sigma_{\eta}^2$	$\kappa_y$	$\rho_2(1)$	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha} + \hat{\beta}$	$\% \hat{\kappa}_y$	$\hat{\kappa}_y$	$\hat{\rho}_2(1)$
0.8250	0.0919	4.0000	0.0893	0.2065 (0.0567)	0.0861 (0.0158)	0.7350 (0.0587)	0.8225	0.0	3.1437	0.1026
0.9590	0.0231	4.0000	0.1059	0.0436 (0.0109)	0.0722 (0.0095)	0.8899 (0.0153)	0.9627	0.0	3.5043	0.1321
0.9750	0.0252	5.0000	0.1614	0.0243 (0.0069)	0.0912 (0.0097)	0.8918 (0.0114)	0.9834	5.20	5.8420	0.2694
0.9000	0.1000	5.0781	0.1486	0.1196 (0.0243)	0.1371 (0.0167)	0.7751 (0.0287)	0.9128	0.0	3.8695	0.2156
0.9165	0.1109	6.0000	0.1775	0.0978 (0.0194)	0.1642 (0.0177)	0.7741 (0.0239)	0.9387	0.8	5.4659	0.3086
0.9798	0.0277	6.0000	0.1944	0.0181 (0.0059)	0.1020 (0.0099)	0.8880 (0.0107)	0.9903	55.20	11.4225	0.4125
0.9400	0.1401	7.0000	0.2334	0.0634 (0.0137)	0.2234 (0.0195)	0.7611 (0.0198)	0.9864	98.80	23.3838	0.7180
0.9856	0.0344	7.0000	0.2529	0.0115 (0.0047)	0.1190 (0.0099)	0.8777 (0.0103)	0.9970	100		0.7339
0.9500	0.1000	8.3666	0.2239	0.0505 (0.0117)	0.1887 (0.0168)	0.7950 (0.0173)	0.9840	94.40	18.0580	0.6100
0.9000	0.2000	8.5955	0.2079	0.1233 (0.0245)	0.2367 (0.0250)	0.7106 (0.0278)	0.9490	60.70	15.0589	0.4943
0.9800	0.0500	10.6041	0.2547	0.0166 (0.0056)	0.1418 (0.0114)	0.8548 (0.0118)	0.9965	100		0.7738
0.9500	0.2000	23.3336	0.2695	0.0549 (0.0132)	0.2642 (0.0198)	0.7328 (0.0201)	0.9970	100		0.9398

Tabla 3.3: Modelos GARCH(1,1) gaussianos estimados para cuatro series generadas por modelos ARSV(1) gaussianos y diagnóstico para las observaciones estandarizadas

	Serie 1	Serie 2	Serie 3	Serie 4
	$\phi = 0.9000$ $\sigma_\eta^2 = 0.1000$	$\phi = 0.9165$ $\sigma_\eta^2 = 0.1109$	$\phi = 0.9856$ $\sigma_\eta^2 = 0.0344$	$\phi = 0.9800$ $\sigma_\eta^2 = 0.0500$
$\kappa_y$ teórica	5.0781	6.0000	7.0000	10.6041
$\rho_2(1)$ teórica	0.1486	0.1775	0.2529	0.2547
$\hat{\omega}$	0.1227 (0.0139)	0.1042 (0.0103)	0.0103 (0.0016)	0.0161 (0.0020)
$\hat{\alpha}$	0.1352 (0.0106)	0.1625 (0.0106)	0.1104 (0.0067)	0.1332 (0.0075)
$\hat{\beta}$	0.7766 (0.0175)	0.7723 (0.0141)	0.8863 (0.0056)	0.8627 (0.0062)
$\hat{\alpha} + \hat{\beta}$	0.9118	0.9348	0.9967	0.9959
$\kappa_y$ implícita	3.8305	5.1604	$\nexists$	$\nexists$
$\rho_2(1)$ implícita	0.2112	0.2962	0.6857	0.7236
$\hat{\varepsilon}_t = \frac{y_t}{\hat{\sigma}_t}$				
Curtosis	4.2122	4.4469	4.1180	4.3241
$Q(20)$	19.07	19.93	17.98	18.88
$r_2(1)$	-0.0080	-0.0080	-0.0080	-0.0060
$r_2(2)$	-0.0130	-0.0150	-0.0070	-0.0080
$r_2(5)$	0.0110	0.0110	0.0190	0.0160
$r_2(10)$	-0.0040	-0.0060	0.0020	0.0000
$Q_2(20)$	18.21	18.28	18.10	17.06

$r_2(k)$ : Autocorrelación de orden  $k$  de las observaciones al cuadrado  $y_t^2$ .

$Q_2(20)$ : Estadístico Box-Ljung para  $y_t^2$  (31.4 es el valor crítico al 5%).



## Capítulo 4

# Observaciones atípicas y heterocedasticidad condicional

En este capítulo se estudian los efectos causados por la presencia de observaciones atípicas en la identificación y estimación de la heterocedasticidad condicional. En primer lugar, se analiza cómo afectan estas observaciones al correlograma de los cuadrados y a los contrastes tradicionales de homocedasticidad condicional, en particular al contraste de McLeod y Li. A continuación, se obtienen los sesgos asintóticos causados por los atípicos en los estimadores de Mínimos Cuadrados Ordinarios (MCO) en procesos ARCH( $p$ ). Los sesgos en muestras finitas, tanto en los estimadores MCO como en los estimadores de Máxima Verosimilitud (MV) de modelos ARCH( $p$ ) y GARCH(1,1), se ilustran mediante estudios de simulación. Se propone un procedimiento para tratar la presencia simultánea de observaciones atípicas y heterocedasticidad condicional en series temporales, y se ilustra en las doce series de rendimientos financieros descritas en el Capítulo 2.

## 4.1 Introducción

Los modelos de heterocedasticidad condicional describen la evolución en el tiempo de las varianzas condicionales. Sin embargo, estos modelos no son siempre capaces de explicar la frecuencia y magnitud de valores extremos que comúnmente aparecen en series temporales; véase, por ejemplo, Bollerslev (1987), Teräsvirta (1996) y Franses y Ghijsels (1999). Las observaciones estandarizadas por la varianza condicional estimada tienen todavía exceso de curtosis, lo que podría explicarse, entre otras posibles causas, por la presencia de observaciones atípicas. La presencia simultánea de heterocedasticidad condicional y observaciones atípicas es un problema difícil de tratar porque dichos fenómenos se confunden fácilmente al causar ambos, exceso de curtosis y, si los atípicos son consecutivos, correlaciones en los cuadrados.

La presencia de observaciones atípicas puede llevar a detectar, de manera espuria, heterocedasticidad condicional y, dado que los métodos tradicionales para la detección de observaciones atípicas se basan en la hipótesis de homocedasticidad, (véase, por ejemplo, Tsay (1988) y Chen y Liu (1993)) si las varianzas condicionales cambian en el tiempo, las observaciones en periodos de alta varianza condicional podrían ser identificadas como atípicas. En este caso, sería la heterocedasticidad condicional la causa de que se detectasen, de manera espuria, observaciones atípicas.

Algunos autores atribuyen la heterocedasticidad condicional encontrada en series temporales a la presencia de atípicos. Por ejemplo, Balke y Fomby (1994) encuentran que las observaciones atípicas son capaces de explicar la mayoría de la no-linealidad observada en varias series temporales de la economía estadounidense. Aggarwal et al. (1999), tras analizar diez mercados financieros, sugieren que muchos de los efectos GARCH encontrados desaparecen una vez que se tienen en cuenta las observaciones más extremas. En este sentido también, Van Dijk et al. (1999) muestran que la presencia de observaciones atípicas pueden llevar al contraste del Multiplicador de

Lagrange, propuesto por Engle (1982), a rechazar la hipótesis de homocedasticidad y proponen un nuevo contraste que es robusto a este tipo de observaciones.

El objetivo de este capítulo es modelizar la presencia simultánea de heterocedasticidad condicional y observaciones atípicas en series temporales y su organización es como sigue. La sección 2 describe los tipos de observaciones atípicas en series temporales. Sus efectos en la identificación y estimación de la heterocedasticidad condicional se analizan en las secciones 3 y 4 respectivamente. En la sección 5 se propone un procedimiento, en tres pasos, para tratar la presencia simultánea de ambos fenómenos. En el primero, se detectan y corrigen los atípicos más extremos con objeto de evitar los posibles sesgos en una primera estimación de las varianzas condicionales. En el segundo, se calculan las observaciones estandarizadas, en las que se detectan y corrigen los atípicos restantes siguiendo los métodos tradicionales para series homocedásticas. En el tercero, se reestiman las varianzas condicionales. En la sección 6 se aplica el procedimiento anterior a las doce series descritas en el Capítulo 2. La sección 7 contiene las conclusiones.

## 4.2 Observaciones atípicas en series temporales

El estudio de observaciones atípicas en series temporales se ha centrado en los modelos lineales ARMA. Fox (1972) propuso considerar dos tipos de observaciones atípicas: atípicos aditivos (AO) y atípicos innovativos (IO). Tsay (1988) amplió a cuatro los tipos de atípicos en modelos ARMA y propuso un procedimiento iterativo para detectarlos. Tolvi (2000) y Peña (2001) contienen revisiones de este tema.

Como se ha definido en la sección 2.4, un modelo  $\text{ARMA}(p, q)$  viene dado por

$$\phi(L)y_t = \theta(L)a_t \quad (4.1)$$

Alternativamente, si el proceso  $y_t$  es estacionario, éste puede expresarse como un pro-

ceso  $AR(\infty)$ :  $\pi(L)y_t = a_t$  donde  $\pi(L) = \phi(L)\theta(L)^{-1}$ , o como un proceso  $MA(\infty)$ :  $y_t = \psi(L)a_t$  donde  $\psi(L) = \theta(L)\phi(L)^{-1}$  siempre que  $y_t$  sea invertible.

Un AO mide un cambio exógeno sobre la serie  $y_t$ . Esto es, en lugar de observar  $y_t$ , observamos la serie  $z_t$ , que está contaminada en el instante  $\tau$  con un atípico de tamaño  $w_A$ , es decir,

$$z_t = y_t + w_A I(t = \tau) = \begin{cases} y_t & \text{si } t \neq \tau \\ y_t + w_A & \text{si } t = \tau \end{cases} \quad (4.2)$$

Un atípico aditivo sólo afecta al nivel de la serie en el instante  $\tau$  y por tanto, el modelo para la serie observada viene dado por  $z_t = w_A I(t = \tau) + \psi(L)a_t$  o, equivalentemente,  $\pi(L)(z_t - w_A I(t = \tau)) = a_t$ .

El IO está generado, posiblemente, por un cambio endógeno en la serie temporal, esto es, la serie observada es, en este caso

$$z_t = \begin{cases} y_t & \text{si } t < \tau \\ y_t + w_I \psi_j & \text{si } t = \tau + j, \quad j > 0 \end{cases} \quad (4.3)$$

donde  $\psi_j$  son los coeficientes de la representación  $MA(\infty)$  correspondiente. Un atípico innovativo afecta a todas las observaciones posteriores al instante  $\tau$  a través de la memoria del proceso ARMA. El modelo para la serie observada es

$$z_t = \psi(L)(w_I I(t = \tau) + a_t)$$

o, equivalentemente,  $\pi(L)z_t = w_I I(t = \tau) + a_t$ .

Las observaciones atípicas afectan a la estructura de correlación de una serie temporal y causan sesgos en los coeficientes de autocorrelación estimados. Dichos sesgos dependen tanto del número de observaciones atípicas presentes en la serie, como de su tamaño y posición; véase, por ejemplo, Chang et al. (1988) y Chan (1995). En particular, un atípico aditivo grande sesga todos los coeficientes de autocorrelación hacia cero. Puesto que los procedimientos tradicionales para la identificación de modelos

ARMA se basan en las autocorrelaciones estimadas, las observaciones atípicas tienen efectos importantes en la identificación del modelo  $\text{ARMA}(p, q)$  correspondiente; véase, por ejemplo, Deutsch et al. (1990). De un modo similar, los atípicos sesgan las estimaciones de los parámetros del modelo ARMA. Los métodos de Mínimos Cuadrados Ordinarios (MCO) y Máxima Verosimilitud (MV) son muy sensibles a la presencia de observaciones atípicas, especialmente a los atípicos aditivos. Por ejemplo, un único AO tiene un efecto enorme en la estimación del parámetro de un modelo  $\text{AR}(1)$ , que tiende hacia cero a medida que el tamaño del atípico tiende a infinito. En el caso de atípicos innovativos, los efectos no son tan fuertes. Este tipo de observaciones atípicas producen un efecto pequeño en la  $\text{fac}$  y, por tanto, en las estimaciones de los parámetros; véase, por ejemplo, Peña (2001) .

Hotta y Tsay (1998) introducen dos tipos de atípicos en modelos GARCH, atípicos en el nivel (LO), que afectan sólo al nivel de la serie y no a la varianza condicional y, atípicos en la volatilidad (VO), que afectan tanto al nivel como a la varianza condicional de la serie.

Considérese, por ejemplo, el proceso incorrelado,  $y_t$ , que sigue un modelo  $\text{GARCH}(1,1)$ . Entonces, AO e IO coinciden y deberíamos distinguir solamente entre LO y VO. Si la serie  $y_t$  está contaminada con un LO, entonces la serie observada,  $z_t$ , viene dada por

$$\begin{aligned} z_t &= y_t + w_L I(t = \tau) \\ \sigma_t^2 &= \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \tag{4.4}$$

Nótese que la varianza condicional de  $z_t$ ,  $\sigma_t^2$ , depende de  $y_{t-1}^2$ , y por tanto no está afectada por el atípico. Sin embargo, si la serie  $y_t$  está contaminada por un atípico en la volatilidad, VO, entonces la serie observada,  $z_t$ , viene dada por

$$\begin{aligned} z_t &= y_t + w_V I(t = \tau) \\ \sigma_t^2 &= \omega + \alpha z_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \tag{4.5}$$

y en este caso, la varianza condicional  $\sigma_t^2$  depende de  $z_{t-1}^2$ , y por tanto está afectada por el atípico.

Si tanto la media como la varianza condicional varían en el tiempo, hay tres tipos posibles de observaciones atípicas: el AO y el IO que puede ser LO o VO. Por ejemplo, considérese el modelo MA(1)-GARCH(1,1) dado por,

$$\begin{aligned} y_t &= (1 - \theta L)a_t \\ a_t &= \varepsilon_t \sigma_t \\ \sigma_t^2 &= \omega + \alpha a_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \tag{4.6}$$

Si ocurre un IO, la serie observada podría ser

$$\begin{aligned} z_t &= (1 - \theta L)\tilde{a}_t \\ \tilde{a}_t &= \varepsilon_t \sigma_t + w_L I(t = \tau) \\ \sigma_t^2 &= \omega + \alpha a_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \tag{4.7}$$

que es IO-LO, o

$$\begin{aligned} z_t &= (1 - \theta L)\tilde{a}_t \\ \tilde{a}_t &= \varepsilon_t \sigma_t + w_V I(t = \tau) \\ \sigma_t^2 &= \omega + \alpha \tilde{a}_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \tag{4.8}$$

que es IO-VO.

## 4.3 Efectos en la identificación de la heterocedasticidad condicional

Los contrastes de homocedasticidad condicional se basan fundamentalmente en las correlaciones muestrales de las observaciones al cuadrado<sup>1</sup> y en contrastes de no normalidad. Por lo tanto, sesgos en el correlograma de las observaciones al cuadrado afectarán a la identificación de la heterocedasticidad condicional. A continuación se muestra cómo las observaciones atípicas en series homocedásticas pueden confundirse con heterocedasticidad condicional y cómo pueden también ocultarla.

### 4.3.1 Contrastes de homocedasticidad condicional

Los contrastes principales para detectar la presencia de heterocedasticidad condicional en series temporales son:

#### El contraste del Multiplicador de Lagrange (LM)

Este contraste fue propuesto por Engle (1982) para detectar efectos ARCH y viene dado por  $TR^2$ , donde  $T$  es el tamaño muestral y  $R^2$  es el coeficiente de determinación de la regresión de las observaciones al cuadrado,  $y_t^2$ , sobre una constante y  $p$  retardos  $y_{t-1}^2, \dots, y_{t-p}^2$ . Bajo la hipótesis nula de homocedasticidad condicional, el estadístico del contraste se distribuye asintóticamente como una variable  $\chi^2$  con  $p$  grados de libertad. Lee (1991) demuestra que este contraste para procesos GARCH( $p, q$ ) es el mismo que para ARCH( $p$ ).

---

<sup>1</sup>Si la serie tiene estructura en la media condicional, los contrastes se aplican a los residuos al cuadrado una vez se ha estimado un modelo para la media.

### Contrastes de portmanteau

Uno de los contrastes de homocedasticidad condicional más usados es el propuesto por McLeod y Li (1983) (McLL), que se basa en el estadístico

$$Q(k) = T(T-2) \sum_{j=1}^k \frac{r^2(j)}{(T-j)}$$

donde  $r(j)$  es la  $j$ -ésima correlación muestral de la serie cuya incorrelación queremos contrastar. Bajo la hipótesis nula de homocedasticidad condicional, McLeod y Li (1983) prueban que la distribución asintótica de  $Q(k)$  aplicado a las observaciones al cuadrado es aproximadamente una  $\chi^2$  con  $k$  grados de libertad, siempre que el momento de orden ocho de la serie exista. Las propiedades de este estadístico en muestras finitas han sido estudiadas por Engle et al. (1985) para modelos GARCH y por Pérez y Ruiz (2002) para modelos de Volatilidad Estocástica.

Recientemente, Peña y Rodriguez (2002) han propuesto un nuevo contraste de portmanteau, que se denotará PR, que mejora significativamente la potencia de los anteriores y que se basa también en las autocorrelaciones de las observaciones al cuadrado. Dicho estadístico viene dado por

$$D(k) = T(1 - |R_k|^{1/k})$$

donde

$$R_k = \begin{pmatrix} 1 & \tilde{r}(1) & \cdots & \tilde{r}(k) \\ \tilde{r}(1) & 1 & \cdots & \tilde{r}(k) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{r}(k) & \tilde{r}(k) & \cdots & 1 \end{pmatrix}$$

y  $\tilde{r}(j)$  es el coeficiente de autocorrelación de orden  $j$  estandarizado, dado por

$$\tilde{r}(j) = \sqrt{\frac{T+2}{T-j}} r(j)$$



Bajo la hipótesis nula de homocedasticidad condicional, Peña y Rodriguez (2002) prueban que este estadístico, aplicado a las observaciones al cuadrado, tiene una distribución asintótica que se puede aproximar por una distribución Gamma,  $\gamma(\alpha, \beta)$ , con  $\alpha = 3k(k + 1)/4(2k + 1)$  y  $\beta = 3k/2(2k + 1)$ .

### Contrastes robustos

Van Dijk et al. (1999) muestran que las rachas de atípicos aditivos hacen que el contraste LM rechace la hipótesis nula de homocedasticidad condicional con más frecuencia que la nominal. Para solucionar este problema, proponen un contraste robusto que parece funcionar adecuadamente a la hora de discriminar entre efectos ARCH verdaderos y los espurios causados por la presencia de atípicos consecutivos. Su idea es estimar el modelo bajo la hipótesis nula usando un estimador robusto. El estadístico que emplean es similar al utilizado en el contraste LM,  $TR^2 \sim \chi_p^2$  pero en este caso  $R^2$  es el coeficiente de determinación de la regresión de  $\psi(r_t)^2$  sobre una constante y  $\psi(r_{t-1})^2, \dots, \psi(r_{t-p})^2$  donde la función  $\psi$ , que aparece representada en la figura 4.1, viene dada por la expresión:

$$\psi(x) = x(1 - H(|x| - c_1))\text{sign}(x) + H(|x| - c_1)(1 - H(|x| - c_2))g(|x|) \quad (4.9)$$

con  $c_1$  y  $c_2$  constantes de ajuste,  $c_1 = 2.576$ ,  $c_2 = 3.291$ ,  $H(x) = I(x > 0)$ ,  $\text{sign}(\cdot)$  es la función signo y  $g(x)$  es un polinomio de orden 5 que hace que la función  $\psi$  sea dos veces continuamente diferenciable. Por otra parte,  $r_t$  tiene la expresión:

$$r_t = \frac{y_t}{\sigma_y \omega_y(y_{t-1})}$$

donde  $\sigma_y$  es un estimador robusto de escala de  $y_t$ , en este caso el MAD (Mediana de las desviaciones absolutas con respecto a la mediana) y  $\omega_y$ , que aparece también representada en la figura 4.1, viene dada por

$$\omega_y(y_{t-1}) = \frac{\psi(d(y_{t-1})^\alpha)}{d(y_{t-1})^\alpha}$$

siendo  $\psi(\cdot)$  la función definida en (4.9) y  $d(\cdot)$  la distancia de Mahalanobis:

$$d(y_{t-1}) = \frac{|y_{t-1} - m_y|}{\sigma_y}$$

donde  $m_y$  es una medida robusta de localización de  $y_{t-1}$ , en concreto, la mediana. La constante  $\alpha$  se toma igual a 2.

Este contraste robusto tiene mejores propiedades que el contraste LM, en cuanto a tamaño y potencia, cuando la proporción de observaciones atípicas es inferior al 1%. Sin embargo, el tamaño y potencia del mismo se deterioran de manera notoria cuando la proporción de observaciones atípicas supera el 5%.

Alternativamente, Lumsdaine y Ng (1999) analizan los efectos de una mala especificación de la media condicional sobre el contraste LM y encuentran que lleva a rechazar la hipótesis nula de homocedasticidad en más casos de los reales. Para solucionar este problema, proponen un contraste robusto que se basa en añadir términos adicionales en el modelo estimado, en particular, proponen añadir retardos de la variable analizada.

### 4.3.2 Sesgos en los correlogramas

En esta sección se ilustran los efectos que la presencia de observaciones atípicas puede tener en el correlograma de las observaciones al cuadrado y consecuentemente, en los contrastes anteriores. A continuación se derivan los sesgos asintóticos causados por observaciones atípicas en el correlograma de los cuadrados de la serie observada.

Sea  $y_t$  la serie de interés, y considérese que la serie observada,  $z_t$ , está contaminada por  $k$  atípicos consecutivos del mismo tamaño,  $w_A$ . En este caso,  $z_t$  viene dada por

$$z_t = \begin{cases} y_t + w_A & \text{si } t = \tau + 1, \tau + 2, \dots, \tau + k \\ y_t & \text{en otro caso.} \end{cases} \quad (4.10)$$

**Proposición 1** *Sea  $y_t$  una serie temporal estacionaria y sea  $z_t$  la serie observada dada*

en (4.10). Entonces

$$\lim_{w_A \rightarrow \infty} r_2(h) = \begin{cases} \frac{k-h-\frac{k^2}{T-h}}{k(1-\frac{k}{T})} & \text{si } 0 \leq h \leq k \\ \frac{-\frac{k^2}{T-h}}{k(1-\frac{k}{T})} & \text{si } h \geq k \end{cases}$$

donde  $r_2(h)$  es la correlación muestral de orden  $h$  de la serie de cuadrados  $z_t^2$ .

**Demostración:** El coeficiente de correlación de orden  $h$  para la serie de cuadrados  $z_t^2$  viene dado por

$$r_2(h) = \frac{\sum_{t=h+1}^T (z_t^2 - \bar{z}^2)(z_{t-h}^2 - \bar{z}^2)}{\sum_{t=h+1}^T (z_t^2 - \bar{z}^2)^2}$$

donde  $\bar{z}^2$  es la media muestral de la serie  $z_t^2$ , que se supone

$$\bar{z}^2 = \frac{1}{T} \sum_{t=1}^T z_t^2 \approx \frac{1}{T-h} \sum_{t=h+1}^T z_t^2.$$

Entonces, cuando  $0 \leq h \leq k$  se tiene que  $r_2(h)$  es igual a

$$\frac{\sum_{t \in \mathbb{T}} y_t^2 y_{t-h}^2 + (y_{\tau+1} + w_A)^2 y_{\tau+1-h}^2 + \dots + (y_{\tau+k} + w_A)^2 (y_{\tau+k-h} + w_A)^2 + \dots + y_{\tau+h+1}^2 (y_{\tau+1} + w_A)^2 - (T-h)\bar{z}^2^2}{\sum_{t \in \mathbb{T}} y_t^4 + (y_{\tau+1} + w_A)^4 + \dots + (y_{\tau+k} + w_A)^4 - T\bar{z}^2^2}$$

donde  $\mathbb{T} = \{h+1, \dots, \tau, \tau+k+1, \dots, T\}$ . Como  $z^2^2 - \frac{k^2}{T^2} w_A^4 + o(w_A^4) \approx \frac{k^2}{(T-h)^2} w_A^4 + o(w_A^4)$ , se tiene que

$$\lim_{w_A \rightarrow \infty} r_2(h) = \lim_{w_A \rightarrow \infty} \frac{(k-h-\frac{k^2}{T-h})w_A^4}{(k-\frac{k^2}{T})w_A^4} = \frac{k-h-\frac{k^2}{T-h}}{k(1-\frac{k}{T})}$$

siempre que  $0 \leq h \leq k$ . Por otra parte, cuando  $h \geq k$  se tiene que  $r_2(h)$  es igual a

$$\frac{\sum_{t \in \mathbb{T}} y_t^2 y_{t-h}^2 + (y_{\tau+1} + w_A)^2 y_{\tau+1-h}^2 + \dots + (y_{\tau+k} + w_A)^2 y_{\tau+k-h}^2 + \dots + y_{\tau+h+1}^2 (y_{\tau+1} + w_A)^2 - (T-h)\bar{z}^2^2}{\sum_{t \in \mathbb{T}} y_t^4 + (y_{\tau+1} + w_A)^4 + \dots + (y_{\tau+k} + w_A)^4 - T\bar{z}^2^2}$$

Y, por tanto,

$$\lim_{w_A \rightarrow \infty} r_2(h) = \lim_{w_A \rightarrow \infty} \frac{(-\frac{k^2}{T-h})w_A^4}{(k-\frac{k^2}{T})w_A^4} = \frac{-\frac{k^2}{T-h}}{k(1-\frac{k}{T})}$$

siempre que  $h \geq k$  ■

Si el tamaño de muestra,  $T$ , es lo suficientemente grande, el límite anterior será muy cercano a  $\frac{k-h}{k}$  para  $h \leq k$  y muy cercano a cero para  $h \geq k$  y por tanto, un único atípico

sesga hacia cero todos los coeficientes de correlación de las observaciones al cuadrado y rachas de atípicos generan correlaciones en los cuadrados de la serie. Nótese que, por ejemplo, tres atípicos consecutivos generarían correlaciones en los cuadrados de orden uno y dos, lo que podría llevar a los contrastes anteriores a detectar de manera espuria la presencia de heterocedasticidad condicional en la serie. Para ilustrar este problema, se han simulado cuatro series de tamaño muestral  $T = 500$ , todas ellas de media cero y varianza uno. La primera,  $a_t$ , está generada mediante un proceso ruido blanco gaussiano. La segunda,  $y_t$ , sigue un proceso GARCH(1,1) con parámetros  $\omega = 0.1$ ,  $\alpha = 0.1$  y  $\beta = 0.8$ , la tercera,  $x_t$ , sigue un modelo EGARCH(1,1) con parámetros  $\omega = -0.001$ ,  $\alpha = 0.07$ ,  $\beta = 0.98$  y  $\gamma = -0.0456$  y la cuarta,  $z_t$ , sigue un proceso ARSV(1) con parámetros  $\phi = 0.98$ ,  $\sigma_\eta^2 = 0.05$  y  $\sigma_\star^2 = 0.5329$ . Todas las series se han contaminado primero con tres atípicos aditivos consecutivos<sup>2</sup> en los instantes  $t = 200, 201$  y  $202$  y segundo, con tres atípicos aislados del mismo signo, en los instantes  $t = 100, 200$  y  $300$ . Si  $S_t$  es la serie original, las series contaminadas se denominan  $S_t^c$  y  $S_t^a$  respectivamente. Todos los atípicos son de tamaño  $w_A$ , igual a 5 desviaciones típicas<sup>3</sup>. La tabla 4.1 muestra los resultados del Monte Carlo para distintos estadísticos descriptivos de las series, en particular para la media, desviación típica y los coeficientes de asimetría, curtosis y correlación de orden uno de las observaciones al cuadrado, basados en 1000 réplicas generadas para cada uno de los procesos descritos anteriormente. El efecto de los atípicos en la desviación típica y en el coeficiente de curtosis es, como se esperaría, el mismo, independientemente de que los atípicos sean consecutivos o aislados y de que la serie original sea ruido blanco, GARCH, EGARCH o ARSV. Dichos coeficientes son mayores en las series contaminadas. El efecto en la media y en el coeficiente de asimetría varía, ya que en el caso de los atípicos consecutivos dos son de signo positivo

<sup>2</sup>Dado que la mayoría de las series financieras son incorreladas y que la presencia de tres atípicos consecutivos del mismo signo generaría una correlación de orden uno en la serie contaminada, se han considerado los dos primeros atípicos con signo positivo y el tercero con signo negativo.

<sup>3</sup>Todas las series han sido generadas con varianza uno, por tanto  $w_A = 5$ .

y uno es negativo mientras que los aislados son todos positivos, pero el tipo de sesgo causado por ambos atípicos en dichos estadísticos es el mismo. Sin embargo, el efecto en el coeficiente de autocorrelación de orden uno de las observaciones al cuadrado,  $r_2(1)$ , depende de si los atípicos son consecutivos o aislados y de si la serie original sigue un proceso ruido blanco, GARCH, EGARCH o ARSV. Las conclusiones más importantes son: (i) atípicos consecutivos aumentan el coeficiente de correlación de orden uno de las observaciones al cuadrado, y por tanto, si ocurren en una serie ruido blanco, pueden llevar a detectar heterocedasticidad condicional espuria y (ii) atípicos aislados disminuyen  $r_2(1)$  y, por tanto, pueden ocultar la heterocedasticidad condicional presente en la serie. Estos efectos se ilustran en la figura 4.2, donde se han representado, como ejemplo, cuatro series generadas con los modelos Normal, GARCH, EGARCH y ARSV anteriores junto con las fac de las observaciones al cuadrado para las series originales y las series contaminadas con tres atípicos consecutivos y aislados. Es claro que la presencia de observaciones atípicas afecta a las correlaciones muestrales de las observaciones al cuadrado. Cuando los atípicos son consecutivos, los dos primeros coeficientes de autocorrelación aumentan significativamente, como se deriva de la Proposición 1, mientras que cuando los atípicos son aislados, coeficientes de autocorrelación significativos pasan a ser no significativos como sucede en los modelos GARCH y EGARCH que se han representado.

La tabla 4.1 contiene además el porcentaje de veces que los contrastes descritos anteriormente rechazan la hipótesis nula de homocedasticidad condicional. Como muestra la tabla, los efectos que la presencia de las observaciones atípicas tiene en los contrastes de homocedasticidad coinciden con lo encontrado por Van Dijk et al. (1999) y Lumsdaine y Ng (1999) con respecto al contraste LM. Nótese que el estadístico LM para contrastar efectos ARCH se ve seriamente afectado. Para la serie  $a_t$ , el contraste LM rechaza la hipótesis nula de homocedasticidad condicional en el 99% de los casos cuando hay tres atípicos consecutivos. Sin embargo, si los atípicos son aislados, su tamaño es

1.90%, mucho menor que el nominal, que es 5%. En los dos contrastes de portmanteau el efecto es el mismo y dichos contrastes se ven también seriamente afectados por la presencia de observaciones atípicas. Como era de esperar, en este caso, el contraste robusto es el menos afectado.

Para las series condicionalmente heterocedásticas, es decir,  $y_t$ ,  $x_t$  y  $z_t$ , los dos contrastes de portmanteau tienen mejores propiedades en términos de potencia y el contraste robusto es el que tiene la menor potencia. Para las misma series, los efectos de las observaciones atípicas dependen de que dichas observaciones sean consecutivas o aisladas. Cuando los atípicos son consecutivos, al igual que lo que sucedía para la serie  $a_t$ , los contrastes van a rechazar la hipótesis nula de homocedasticidad condicional en más del 99% de los casos, excepto el contraste robusto que mantiene su potencia que a pesar de ser muy baja, no se ve afectada por la presencia de atípicos. Por otra parte, si los atípicos son aislados, los contrastes van a rechazar la hipótesis nula de homocedasticidad condicional en menos casos de los que debería. Por tanto, es claro que atípicos aislados pueden ocultar heterocedasticidad condicional genuina.

Nótese que para los modelos EGARCH, los cuatro contrastes tienen poca potencia. Cuando la serie no tiene atípicos, el contraste LM rechaza la hipótesis nula de homocedasticidad condicional sólo en el 26.40% de los casos cuando la alternativa es en realidad la cierta. En este caso, el contraste McLL es el que tiene más potencia, rechazando la hipótesis nula en un 61% de las veces.

Aunque la presencia de atípicos consecutivos en una serie temporal pueda parecer poco realista, ésta es una situación bastante común en la práctica. Por ejemplo, en un modelo AR(2), si la serie está contaminada con un sólo atípico, los residuos están contaminados con tres atípicos consecutivos lo que podría llevar a los contrastes anteriores a detectar heterocedasticidad condicional cuando en realidad se tendría una serie ruido blanco con atípicos.

## 4.4 Efectos en la estimación de la heterocedasticidad condicional

En esta sección se analizan los efectos que las observaciones atípicas tienen en la estimación de los modelos ARCH y GARCH. Este problema ha sido previamente estudiado en la literatura con resultados contradictorios. Por ejemplo, Mendes (2000) establece las propiedades de robustez de los estimadores de MV de los modelos ARCH, obteniendo su curva de influencia y punto de ruptura. Mediante experimentos de Monte Carlo ilustra los sesgos causados por la presencia de observaciones atípicas en los estimadores de cuasi-máxima verosimilitud (QMV) de los modelos GARCH(1,1) concluyendo que el estimador del parámetro  $\alpha$  está sesgado hacia cero mientras que el del parámetro  $\beta$  está sesgado hacia uno y además, los efectos son menores cuanto mayor es el tamaño muestral. Verhoeven y McAleer (2000) y Gregory y Reeves (2001) llegan a resultados contrarios. Mediante un experimento de Monte Carlo, Gregory y Reeves (2001) concluyen que cuando hay atípicos, el parámetro  $\alpha$  se sobreestima y el parámetro  $\beta$  se infraestima. Verhoeven y McAleer (2000) llegan a los mismos resultados analizando cinco series diarias de rendimientos financieros.

### 4.4.1 Estimación de modelos ARCH

Sea  $y_t$  una serie generada por un proceso ARCH( $p$ ),

$$\begin{aligned} y_t &= \varepsilon_t \sigma_t \\ \sigma_t^2 &= \omega + \alpha_1 y_{t-1}^2 + \dots + \alpha_p y_{t-p}^2 \end{aligned} \tag{4.11}$$

donde  $\varepsilon_t$  se supone i.i.d. con media cero y varianza uno,  $\omega \geq 0$ ,  $0 < \alpha_i < 1$ ,  $\forall i$  y  $\sum_{i=1}^p \alpha_i < 1$ .

### Mínimos Cuadrados Ordinarios

Considérese la representación del modelo ARCH( $p$ ) como un modelo AR( $p$ ) para  $y_t^2$  como sigue

$$y_t^2 = \omega + \sum_{i=1}^p \alpha_i y_{t-i}^2 + v_t \quad t = p+1, \dots, T \quad (4.12)$$

donde  $v_t = y_t^2 - \sigma_t^2$  es un proceso incorrelado, de media cero y varianza constante pero condicionalmente heterocedástico.

Equivalentemente,

$$\mathbf{Y} = \mathbf{X}\theta + \mathbf{V} \quad (4.13)$$

donde

$$\mathbf{Y} = \begin{bmatrix} y_{p+1}^2 \\ y_{p+2}^2 \\ \vdots \\ y_T^2 \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & y_p^2 & y_{p-1}^2 & \cdots & y_1^2 \\ 1 & y_{p+1}^2 & y_p^2 & \cdots & y_2^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & y_{T-1}^2 & y_{T-2}^2 & \cdots & y_{T-p}^2 \end{bmatrix} \quad \mathbf{V} = \begin{bmatrix} v_{p+1} \\ v_{p+2} \\ \vdots \\ v_T \end{bmatrix}$$

y  $\theta$  es el vector de parámetros definido como  $\theta = [\omega \quad \alpha]'$  y  $\alpha = [\alpha_1 \dots \alpha_p]'$ . Entonces, el estimador de mínimos cuadrados ordinarios (MCO) de  $\theta$  viene dado por

$$\hat{\theta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}. \quad (4.14)$$

Si el momento de cuarto orden de  $y_t$  existe,  $\hat{\theta}$  es consistente, ya que  $E(\mathbf{X}'\mathbf{V}) = 0$ ; ver, por ejemplo Poterba y Summers (1986). Aunque el estimador MCO es el mejor estimador lineal insesgado, es ineficiente debido a la dependencia cuadrática del proceso  $v_t$ .

### Máxima Verosimilitud

En el modelo (4.11),  $\sigma_t^2$  es la varianza de  $y_t$  condicional a  $Y_{t-1} = \{y_{t-1}, y_{t-2}, \dots, y_1\}$ . Si además  $\varepsilon_t$  es gaussiano, la distribución de  $y_t | Y_{t-1}$  es  $N(0, \sigma_t^2)$  y consecuentemente,



la estimación de estos modelos<sup>4</sup> es inmediata maximizando la función de verosimilitud dada por

$$f(Y_t; \omega, \alpha_i) = \prod_{t=p+1}^T \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp \left\{ -\frac{y_t^2}{2\sigma_t^2} \right\} \quad (4.15)$$

Maximizar la función de verosimilitud es equivalente a maximizar su logaritmo, que viene dado por

$$L = -\frac{T-p}{2} \log(2\pi) + \sum_{t=p+1}^T l_t \quad (4.16)$$

donde

$$l_t = -\frac{1}{2} \left( \log \sigma_t^2 + \frac{y_t^2}{\sigma_t^2} \right)$$

Por otra parte, se tiene que

$$\sigma_t^2 = x_t' \theta = \omega + \sum_{i=1}^p \alpha_i y_{t-i}^2$$

con

$$x_t' = [1 \quad y_{t-1}^2 \dots y_{t-p}^2]$$

Las derivadas de primer y segundo orden de  $l_t$  con respecto a  $\theta$  vienen dadas por

$$\frac{\partial l_t}{\partial \theta} = -\frac{1}{2\sigma_t^2} \left( 1 - \frac{y_t^2}{\sigma_t^2} \right) x_t \quad \frac{\partial^2 l_t}{\partial \theta \partial \theta'} = \frac{1}{2\sigma_t^4} \left( 1 - \frac{2y_t^2}{\sigma_t^2} \right) x_t x_t'$$

Por lo tanto,

$$\frac{\partial L}{\partial \theta} = \mathbf{X}' \mathbf{Z}^*$$

donde  $\mathbf{X}$  es la matriz de dimensión  $(p+1) \times (T-p)$  cuyas columnas son los vectores  $x_t$  y  $\mathbf{Z}^*$  es un vector de dimensión  $(T-p) \times 1$  cuyos elementos son:

$$z_t^* = -\frac{1}{2\sigma_t^2} \left( 1 - \frac{y_t^2}{\sigma_t^2} \right)$$

y

$$\frac{\partial^2 L}{\partial \theta \partial \theta'} = \mathbf{X}' \mathbf{W} \mathbf{X}$$

---

<sup>4</sup>La estimación MV de modelos GARCH( $p, q$ ) es análoga.

donde  $\mathbf{W}$  es una matriz diagonal de pesos:

$$w_t = -\frac{1}{2\sigma_t^4} \left( 1 - \frac{2y_t^2}{\sigma_t^2} \right)$$

de dimensión  $(T-p) \times (T-p)$ . El estimador MV de  $\theta$  es, por tanto, la solución de la siguiente ecuación

$$\frac{\partial L}{\partial \theta} = \mathbf{X}'\mathbf{Z}^* = 0$$

La maximización del logaritmo de la función de verosimilitud se lleva a cabo, usualmente, mediante técnicas de optimización. Engle (1982) y Bollerslev (1986) describen críticamente algunos procedimientos iterativos alternativos que han sido aplicados con éxito en la estimación de los modelos ARCH.

#### 4.4.2 Sesgos en los estimadores MCO de los modelos ARCH( $p$ )

En esta sección se estudian los efectos causados por atípicos en el nivel sobre los estimadores MCO de un proceso ARCH( $p$ ) definido en (4.11). En primer lugar, se supone que dicho proceso está contaminado en el instante  $\tau$  con un atípico en el nivel de tamaño  $w_A$ . Y a continuación, se analizan los efectos cuando la serie está contaminada con  $k$  atípicos consecutivos.

##### Un único atípico

La siguiente Proposición proporciona una expresión para los sesgos asintóticos causados por un único atípico en los estimadores MCO del modelo ARCH( $p$ )

**Proposición 2** *Sea  $y_t$  un proceso ARCH( $p$ ) dado por (4.11). Si la serie observada es  $z_t$  dada por (4.2), los estimadores MCO de los parámetros  $\omega$  y  $\alpha = [\alpha_1, \dots, \alpha_p]'$  verifican que*

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \infty \\ -\frac{1}{T-2p} \mathbf{1} \end{bmatrix} \quad \text{donde } \mathbf{1} = [1 \quad 1 \quad \dots \quad 1]'. \quad (4.17)$$

**Demostración:** Como se muestra en (4.14), los estimadores MCO de  $\omega, \alpha_1, \dots, \alpha_p$  vienen dados por

$$\begin{bmatrix} \hat{\omega} \\ \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_p \end{bmatrix} = \begin{bmatrix} T-p & \sum_{t=p}^{T-1} z_t^2 & \cdots & \sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2 \\ \sum_{t=p}^{T-1} z_t^2 & \sum_{t=p}^{T-1} z_t^4 & \cdots & \sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2 & \sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2 & \cdots & \sum_{t=p}^{T-1} z_t^4 z_{t-p+1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=p+1}^T z_t^2 \\ \sum_{t=p+1}^T z_t^2 z_{t-1}^2 \\ \vdots \\ \sum_{t=p+1}^T z_t^2 z_{t-p}^2 \end{bmatrix} \quad (4.18)$$

Como  $\sum_{t=p}^{T-1} z_t^4 = w_A^4 + o(w_A^4)$ , al igual que  $\sum_{t=p}^{T-1} z_{t-1}^4, \dots, \sum_{t=p}^{T-1} z_{t-p+1}^4$  y  $\sum_{t=p}^{T-1} z_t^2 = w_A^2 + o(w_A^2)$ , al igual que  $\sum_{t=p}^{T-1} z_{t-1}^2, \dots, \sum_{t=p}^{T-1} z_{t-p+1}^2$  y que  $\sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2, \dots, \sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2$ , entonces se tiene,

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} = \lim_{w_A \rightarrow \infty} \begin{bmatrix} T-p & (w_A^2 + o(w_A^2))\mathbf{1}' \\ (w_A^2 + o(w_A^2))\mathbf{1} & (w_A^2 + o(w_A^2))\mathbf{F} \end{bmatrix}^{-1} \begin{bmatrix} w_A^2 + o(w_A^2) \\ (w_A^2 + o(w_A^2))\mathbf{1} \end{bmatrix}$$

donde  $\mathbf{F} = \{f_{ij}\}$  es una matriz simétrica con  $f_{ij} = 1$  para  $i = 1, \dots, p, j = i+1, \dots, p$  y  $f_{ii} = w_A^2$  para  $i = 1, \dots, p$ . Así,

$$\begin{aligned} \lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} &= \lim_{w_A \rightarrow \infty} \frac{1}{(T-2p)w_A^{4p}} \begin{bmatrix} w_A^{4p} + o(w_A^{4p}) & (w_A^{4p-2} + o(w_A^{4p-2}))\mathbf{1}' \\ (-w_A^{4p-2} + o(w_A^{4p-2}))\mathbf{1} & \mathbf{V} \end{bmatrix} \begin{bmatrix} w_A^2 + o(w_A^2) \\ (w_A^2 + o(w_A^2))\mathbf{1} \end{bmatrix} \\ &= \lim_{w_A \rightarrow \infty} \frac{1}{(T-2p)w_A^{4p}} \begin{bmatrix} w_A^{4p+2} + o(w_A^{4p+2}) \\ (-w_A^{4p} + o(w_A^{4p}))\mathbf{1} \end{bmatrix} = \begin{bmatrix} \infty \\ -\frac{1}{T-2p}\mathbf{1} \end{bmatrix} \end{aligned}$$

donde  $\mathbf{V} = \{v_{ij}\}$  es una matriz simétrica con  $v_{ij} = o(w_A^{4p-2})$ . Por lo tanto, para un modelo ARCH( $p$ ) contaminado con un atípico de tamaño  $w_A$ , se tiene que los estimadores MCO de  $\omega$  y  $\alpha = [\alpha_1 \dots \alpha_p]'$  verifican

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \infty \\ -\frac{1}{T-2p}\mathbf{1} \end{bmatrix} \blacksquare$$

Se puede observar que, como era de esperar, la presencia de un atípico cuyo tamaño tiende a infinito en una serie generada por un modelo ARCH( $p$ ) hace que la varianza

marginal estimada,  $\hat{\omega}/(1 - \sum_{i=1}^p \hat{\alpha}_i)$ , tienda a infinito y la dependencia dinámica de la varianza condicional desaparezca.

Además, dado que la persistencia en un modelo ARCH( $p$ ) puede medirse mediante  $\sum_{i=1}^p \alpha_i$ , véase, por ejemplo, Bollerslev (1986), ésta tiende hacia  $-\frac{p}{T-2p}$  cuando  $w_A$  tiende a infinito. Nótese también que en este caso, cuando el tamaño muestral es grande, la persistencia es muy cercana a cero.

Como ilustración, la figura 4.3 representa los cuantiles de los estimadores MCO, basados en 500 réplicas, de modelos ARCH(1) y ARCH(4) con parámetros  $\omega = 0.8$  y  $\alpha = 0.2$  y  $\omega = 0.55$ ,  $\alpha_1 = 0.2$ ,  $\alpha_2 = 0.1$ ,  $\alpha_3 = 0.1$  y  $\alpha_4 = 0.05$  respectivamente. Ambos modelos han sido contaminados con un único atípico. Como se puede observar en el gráfico, a medida que el tamaño del atípico aumenta, los estimadores de los parámetros tienden a los valores límite dados en la Proposición 2. Para ambos modelos  $\omega$  crece indefinidamente con  $w_A$ . En cuanto a los parámetros  $\alpha_i$ , obsérvese que para un tamaño  $w_A$  suficientemente grande, éste se estima como  $-\frac{1}{T-2}$  en el caso del ARCH(1), que es  $-0.0020$ ,  $-0.0010$  y  $-0.0002$  para  $T = 500, 1000$  y  $5000$  respectivamente. Para el ARCH(4), los estimadores de  $\alpha_i$  tienden a  $-\frac{1}{T-8}$ , para  $i = 1, 2, 3, 4$ . Nótese también que, como era de esperar, los sesgos causados por un atípico de tamaño  $w_A$  fijo y finito, son menores a medida que el tamaño muestral,  $T$ , aumenta. Así, por ejemplo, para el modelo ARCH(1) con un atípico de tamaño 7 desviaciones típicas, la mediana de los estimadores MCO de  $\omega$  y  $\alpha$  es 1.03 y 0.05 respectivamente cuando  $T = 500$ , 0.95 y 0.09 cuando  $T = 1000$  y 0.84 y 0.16 cuando  $T = 5000$ . Otro hecho importante es que, cuando en una serie existen atípicos muy grandes, éstos pueden llevar a que los parámetros estimados no satisfagan las condiciones de no negatividad necesarias para que el modelo esté bien definido.

### $k$ atípicos consecutivos

Considérese ahora que la serie está contaminada por  $k$  atípicos consecutivos del mismo tamaño,  $w_A$ . En este caso, la serie observada,  $z_t$ , viene dada por (4.10). Como se verá a continuación, los sesgos serán diferentes dependiendo de la relación entre el número de atípicos y el orden del modelo,  $k$  y  $p$  respectivamente.

**Proposición 3** *Sea  $y_t$  un proceso ARCH(1) dado por (4.11) con  $p = 1$ . Si la serie observada es  $z_t$  en (4.10) con  $k \geq 1^5$ , los estimadores MCO de los parámetros  $\omega$  y  $\alpha_1$  verifican que*

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha}_1 \end{bmatrix} = \begin{bmatrix} \infty \\ \frac{(T-1)(k-1)-k^2}{(T-1)k-k^2} \end{bmatrix}. \quad (4.19)$$

**Demostración:** Dado que los estimadores MCO de los parámetros  $\omega$  y  $\alpha_1$  vienen dados por (4.14), sustituyendo en dicha expresión  $z_t$  y teniendo en cuenta que  $\sum_{t=1}^{T-1} z_t^2 = kw_A^2 + o(w_A^2)$  y que  $\sum_{t=1}^{T-1} z_t^4 = kw_A^4 + o(w_A^4)$ , se tiene que

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha}_1 \end{bmatrix} = \lim_{w_A \rightarrow \infty} \frac{1}{(T-1)kw_A^4 - (kw_A^2)^2} \begin{bmatrix} kw_A^4 + o(w_A^4) & -kw_A^2 + o(w_A^2) \\ -kw_A^2 + o(w_A^2) & T-1 \end{bmatrix} \begin{bmatrix} kw_A^2 + o(w_A^2) \\ (k-1)w_A^4 + o(w_A^4) \end{bmatrix}$$

Y, por tanto

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha}_1 \end{bmatrix} = \begin{bmatrix} \infty \\ \frac{(T-1)(k-1)-k^2}{(T-1)k-k^2} \end{bmatrix} \blacksquare$$

En este caso, la varianza marginal estimada tiende a infinito, como en el caso de un único atípico. Sin embargo, la estimación del parámetro  $\alpha_1$  tiende a un valor próximo a uno cuando  $k$  es lo suficientemente grande. Nótese que en un modelo ARCH(1), el parámetro  $\alpha_1$  mide, tanto la dependencia dinámica de la varianza como la persistencia. Por lo tanto, cuando aparecen un número grande,  $k$ , de atípicos consecutivos en una serie generada por un modelo ARCH(1), se podría concluir erróneamente que existe una raíz unitaria en los segundos momentos.

<sup>5</sup>Nótese que si  $k = 1 = p$  es el caso de un único atípico y el resultado coincide con el de la Proposición 2.

Como ilustración de este caso, la figura 4.4 representa los cuantiles de los estimadores MCO, basados en 500 réplicas, del modelo ARCH(1) considerado anteriormente, contaminado en este caso con 2 y 20 atípicos consecutivos. Como se puede observar en el gráfico, a medida que el tamaño del atípico aumenta, los estimadores de los parámetros tienden a los valores límite dados en la Proposición 3. Nótese que  $\omega$  crece indefinidamente con  $w_A$ , mientras que  $\alpha_1$  se estima como  $\frac{T-5}{2T-6}$  en el caso de dos atípicos, que es aproximadamente 0.5 para los tamaños de muestra considerados y  $\frac{19T-419}{20T-420}$ , en el caso de veinte atípicos, que es aproximadamente 0.95. Al igual que cuando hay un único atípico, los sesgos van a depender en gran medida del tamaño del atípico,  $w_A$ , y del tamaño muestral  $T$ . Por ejemplo, en el caso que ilustra la figura 4.4, un modelo ARCH(1) contaminado con dos atípicos consecutivos y de tamaño muestral  $T = 500$ , la mediana de los estimadores de  $\omega$  y  $\alpha_1$  es 0.06 y 0.47 respectivamente cuando  $w_A = 5$ , 0.11 y 0.48 cuando  $w_A = 7$  y 0.21 y 0.49 cuando  $w_A = 10$ .

A continuación se va a estudiar cuál es el efecto de  $k$  atípicos consecutivos en un modelo ARCH( $p$ ) con  $p > 1$ .

**Proposición 4** *Sea  $y_t$  un proceso ARCH( $p$ ) dado por (4.11). Si la serie observada es  $z_t$  en (4.10) con  $k \geq p > 1$ , se tiene que los estimadores MCO de los parámetros  $\omega$  y  $\alpha = [\alpha_1 \dots \alpha_p]'$  verifican que*

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \infty \\ a \end{bmatrix} \quad (4.20)$$

donde

$$a = \left[ \frac{2k^2 + (2k - p)(T - p)}{2k^2 + (2k - p + 1)(T - p)} \quad 0 \quad \dots \quad 0 \quad \frac{(T - p)}{2k^2 + (2k - p + 1)(T - p)} \right]'$$

**Demostración:** Los estimadores MCO de los parámetros  $\omega$  y  $\alpha = [\alpha_1 \dots \alpha_p]'$  vienen dados por (4.18). Por tanto, sustituyendo en dicha expresión  $z_t$ , y teniendo en cuenta que  $\sum_{t=p}^{T-1} z_t^2 = kw_A^2 + o(w_A^2)$ , al igual que  $\sum_{t=p}^{T-1} z_{t-1}^2, \dots, \sum_{t=p}^{T-1} z_{t-p+1}^2$  y que  $\sum_{t=p}^{T-1} z_t^4 -$

$kw_A^4 + o(w_A^4)$ , como  $\sum_{t=p}^{T-1} z_{t-1}^4, \dots, \sum_{t=p}^{T-1} z_{t-p+1}^4$  y que  $\sum_{t=p}^{T-1} z_t^2 z_{t+1}^2 = (k-1)w_A^4 + o(w_A^4), \dots, \sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2 = (k-p+1)w_A^4 + o(w_A^4)$ , se tiene que

$$\begin{aligned} \lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} &= \lim_{w_A \rightarrow \infty} \begin{bmatrix} T & p & (kw_A^2 + o(w_A^2))\mathbf{1}' \\ (kw_A^2 + o(w_A^2))\mathbf{1} & (w_A^4 + o(w_A^4))\mathbf{M} & \end{bmatrix}^{-1} \begin{bmatrix} kw_A^2 + o(w_A^2) \\ (w_A^4 + o(w_A^4))\mathbf{B} \end{bmatrix} \\ &= \lim_{w_A \rightarrow \infty} \begin{bmatrix} c_1 & c_2 \ell' \\ c_2 \ell & (w_A^4 + o(w_A^4))\mathbf{D} \end{bmatrix} \begin{bmatrix} kw_A^2 + o(w_A^2) \\ (w_A^4 + o(w_A^4))\mathbf{B} \end{bmatrix} \end{aligned}$$

donde  $\mathbf{1} = [1 \ 1 \ \dots \ 1]'$ .  $\mathbf{M} = \{m_{ij}\}$  es una matriz simétrica, con  $m_{ij} = k + i - j$  para  $i = 1, \dots, p$  y  $j = i, \dots, p$ .  $\mathbf{B} = \{b_i\}$  con  $b_i = k - i$  para  $i = 1, \dots, p$ .  $\mathbf{D} = \{d_{ij}\}$  es una matriz, también simétrica, con  $d_{11} = d_{pp} = \frac{1}{2w_A^4} \frac{-2k^2 + (2k-p+2)(T-p)}{-2k^2 + (2k-p+1)(T-p)}$ ,  $d_{ii} = \frac{1}{w_A^4}$  para  $i = 2, \dots, p-1$ ,  $d_{ii+1} = \frac{1}{2w_A^4}$  para  $i = 2, \dots, p-1$ ,  $d_{1p} = \frac{1}{2w_A^4} \frac{T-p}{-2k^2 + (2k-p+1)(T-p)}$  y  $d_{ij} = 0$  en otro caso.  $\ell = [1 \ 0 \ 0 \ \dots \ 0 \ 1]'$ ,  $c_1 = \frac{2^{p-1}[k-(p-1)/2]}{2^{p-2}[-2k^2 + (2k-p+1)(T-p)]}$  y  $c_2 = \frac{1}{w_A^4} \frac{k}{-2k^2 + (2k-p+1)(T-p)}$ . Por lo tanto,

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \infty \\ a \end{bmatrix} \blacksquare$$

Al igual que en los casos anteriores, la varianza marginal estimada tiende a infinito con  $w_A$ , el tamaño del atípico. En cuanto a los estimadores de los parámetros  $\alpha_i$ , nótese que solamente  $\alpha_1$  y  $\alpha_p$  tienden a cantidades distintas de cero, siendo el límite del estimador de  $\alpha_1$  muy próximo a uno para un número suficientemente grande de atípicos consecutivos. Nótese además que la persistencia,  $\sum_{i=1}^p \alpha_i$ , tiende a  $\frac{2k^2 + (2k-p-1)(T-p)}{2k^2 + (2k-p+1)(T-p)}$  que, como era de esperar, coincide con el valor obtenido en la Proposición 3 para el caso particular  $k = 1 = p$ .

Como ilustración de estos efectos, la figura 4.5 contiene los cuantiles de los estimadores MCO, basados en 500 réplicas, del modelo ARCH(4) anterior, contaminado en este caso con 20 atípicos consecutivos. Como se puede observar en el gráfico, a medida que el tamaño del atípico aumenta, los estimadores de los parámetros tienden a los valores límite dados en la Proposición 4. Análogo a los casos anteriores  $\omega$  crece

indefinidamente con  $w_A$ , mientras que los parámetros  $\alpha_1, \alpha_2, \alpha_3$  y  $\alpha_4$ , se estiman como

$\frac{36T-944}{37T-948}, 0, 0$  y  $\frac{-T+4}{37T-948}$  respectivamente.

Finalmente, se va analizar el efecto de un número de atípicos consecutivos menor que el orden del proceso ARCH( $p$ ).

**Proposición 5** *Sea  $y_t$  un proceso ARCH( $p$ ) dado por (4.11) con  $p > 2$ . Si la serie observada es  $z_t$  en (4.10) con  $k = 2$ , se tiene que los estimadores MCO de los parámetros  $\omega$  y  $\alpha = [\alpha_1 \dots \alpha_p]'$  verifican que*

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} \approx \begin{cases} \begin{bmatrix} \infty \\ a_{par} \end{bmatrix} & \text{si } p \text{ es par} \\ \begin{bmatrix} \infty \\ a_{impar} \end{bmatrix} & \text{si } p \text{ es impar} \end{cases} \quad (4.21)$$

donde

$$a_{par} = \left[ \frac{p}{p+1} + \frac{p^2}{T-p(p+3)} \quad -\frac{p}{p+1} + \frac{2p}{T-p(p+3)} \quad \dots \quad \frac{1}{p+1} + \frac{p^2}{T-p(p+3)} \right]'$$

$$a_{impar} = \left[ \frac{p}{p+1} + \frac{(p+1)^2}{T-p(p+1)^2} \quad \frac{p}{p+1} \quad \dots \quad \frac{1}{p+1} + \frac{(p+1)^2}{T-p(p+1)^2} \right]'$$

**Demostración:** Los estimadores MCO de los parámetros  $\omega$  y  $\alpha = [\alpha_1 \dots \alpha_p]'$  vienen dados por (4.18). Por tanto, sustituyendo en dicha expresión  $z_t$ ,  $k = 2$  y teniendo en cuenta los órdenes de los elementos como, por ejemplo,  $\sum_{t=p}^T z_t^2 = 2w_A^2 + o(w_A^2)$  que es del mismo orden que  $\sum_{t=p}^{T-1} z_{t-1}^2, \dots, \sum_{t=p}^{T-1} z_{t-p+1}^2$  y que  $\sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2 = w_A^2 + o(w_A^2)$  y  $\sum_{t=p}^{T-1} z_t^2 z_{t+1}^2 = w_A^4 + o(w_A^4)$ , se tiene que

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} = \lim_{w_A \rightarrow \infty} \begin{bmatrix} T-p & (2w_A^2 + o(w_A^2))\mathbf{1}' \\ (2w_A^2 + o(w_A^2))\mathbf{1} & (w_A^4 + o(w_A^4))\mathbf{R} \end{bmatrix}^{-1} \begin{bmatrix} 2w_A^2 + o(w_A^2) \\ (w_A^4 + o(w_A^4))\mathbf{S} \end{bmatrix}$$



$$= \lim_{w_A \rightarrow \infty} \frac{1}{w_A^2} \begin{bmatrix} A_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}^{-1} \begin{bmatrix} 2w_A^2 + o(w_A^2) \\ (w_A^4 + o(w_A^4))\mathbf{S} \end{bmatrix}$$

donde  $\mathbf{1} = [1 \ 1 \ \dots \ 1]'$ ,  $\mathbf{R} = \{r_{ij}\}$  es una matriz simétrica, con  $r_{ii} = 2$ ,  $r_{ii+1} = 1$  para  $i = 1, \dots, p-1$ ,  $r_{ij} = \frac{e_{ij}}{w_A^2}$  para  $j = i+2, \dots, p$  y  $e_{ij}$  son constantes reales.  $\mathbf{S} = \{s_i\}$  donde  $s_1 = 1$ ,  $s_i = \frac{ss_i}{w_A^4}$  for  $i = 2, \dots, p$  y  $ss_i$  son constantes reales. Por lo tanto,

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} = \lim_{w_A \rightarrow \infty} \begin{bmatrix} X & -X\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}X & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}X\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \end{bmatrix} \begin{bmatrix} 2 + o(1) \\ (w_A^2 + o(w_A^2))\mathbf{S} \end{bmatrix}$$

donde  $X = (A_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}$ ,  $\mathbf{A}_{22} = w_A^2\tilde{\mathbf{R}}$  puede aproximarse por  $w_A^2\tilde{\mathbf{R}}$  donde  $\tilde{\mathbf{R}} = \{\tilde{r}_{ij}\}$  es una matriz de Toeplitz, simétrica con  $\tilde{r}_{ii} = 2$ ,  $\tilde{r}_{ii+1} = 1$  para  $i = 1, \dots, p$ ,  $\tilde{r}_{ij} = 0$  para  $j = i+2, \dots, p$ . Entonces,  $\mathbf{A}_{22}^{-1} \approx \frac{1}{w_A^2}\tilde{\mathbf{R}}^{-1} = \frac{1}{w_A^2}\mathbf{T}$  y  $\mathbf{T} = \{t_{ij}\}$  con  $t_{ij} = (-1)^{i+j}\frac{(p+1-j)^i}{p+1}$  para  $i = 1, \dots, p$  y  $j = i, \dots, p$ . Después de algunos cálculos se obtiene

$$X \approx \begin{cases} \frac{w_A^2}{T-p(p+3)} & \text{si } p \text{ es par} \\ \frac{w_A^2}{T-p-(p+1)^2} & \text{si } p \text{ es impar} \end{cases}$$

$$-X\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \approx \begin{cases} -\frac{1}{T-p(p+3)} \begin{bmatrix} p & 2 & p-2 & 4 & p-4 & \dots & p \end{bmatrix} & \text{si } p \text{ es par} \\ -\frac{1}{T-p-(p+1)^2} \begin{bmatrix} p+1 & 0 & p+1 & 0 & p+1 & \dots & p+1 \end{bmatrix} & \text{si } p \text{ es impar} \end{cases}$$

$$-\mathbf{A}_{22}^{-1}\mathbf{A}_{21}X = (-X\mathbf{A}_{12}\mathbf{A}_{22}^{-1})'$$

Por lo tanto,

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} \approx \begin{cases} \begin{bmatrix} \infty \\ a_{\text{par}} \end{bmatrix} & \text{si } p \text{ es par} \\ \begin{bmatrix} \infty \\ a_{\text{impar}} \end{bmatrix} & \text{si } p \text{ es impar} \end{cases}$$

donde

$$a_{\text{par}} = \left[ \frac{p}{p+1} + \frac{p^2}{T - p(p+3)} \quad \frac{p}{p+1} + \frac{2p}{T - p(p+3)} \quad \cdots \quad \frac{1}{p+1} + \frac{p^2}{T - p(p+3)} \right]',$$

$$a_{\text{impar}} = \left[ \frac{p}{p+1} + \frac{(p+1)^2}{T - p(p+1)^2} \quad -\frac{p}{p+1} \quad \cdots \quad \frac{1}{p+1} + \frac{(p+1)^2}{T - p(p+1)^2} \right]' \blacksquare$$

Como en los casos anteriores, la varianza marginal estimada tiende a infinito con el tamaño del atípico. La persistencia estimada,  $\sum \hat{\alpha}_i$ , tiende a un valor muy próximo a 0.5 cuando el atípico tiende a infinito y el tamaño muestral es lo suficientemente grande. Más concretamente, la aproximación dada por la Proposición 5 para el límite de la persistencia cuando hay dos atípicos consecutivos sería,  $\frac{p}{2(p+1)}$  si  $p$  es par y 0.5 si  $p$  es impar.

La figura 4.5 contiene los cuantiles de los estimadores MCO, basados en 500 réplicas, del modelo ARCH(4) anterior, contaminado en este caso con 2 atípicos consecutivos. En este caso, la Proposición 5 no aporta los límites exactos de los estimadores sino valores aproximados de dichos límites. Como se puede observar en el gráfico,  $\omega$  crece indefinidamente con  $w_A$ . Los estimadores de los parámetros  $\alpha_1, \alpha_2, \alpha_3$  y  $\alpha_4$ , se aproximarían, siguiendo la Proposición 5, por  $\frac{4}{5} + \frac{16}{T-28}$ ,  $\frac{3}{5} + \frac{8}{T-28}$ ,  $\frac{2}{5} + \frac{4}{T-28}$  y  $\frac{1}{5} + \frac{16}{T-28}$  respectivamente.

La siguiente Proposición se ha obtenido extrapolando los resultados obtenidos, usando el programa *Maple V*, para los distintos casos particulares que se muestran en la Tabla 4.2.

**Proposición 6** *Sea  $y_t$  un proceso ARCH( $p$ ) dado por (4.11). Si la serie observada es  $z_t$  en (4.10) con  $k = p - 1$ , se tiene que los estimadores MCO de los parámetros  $\omega$ , y  $\alpha = [\alpha_1 \dots \alpha_p]'$  verifican que*

$$\lim_{w_A \rightarrow \infty} \begin{bmatrix} \hat{\omega} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \infty \\ a \end{bmatrix} \quad (4.22)$$

donde

$$a = \begin{bmatrix} \frac{(2p-3)T - (6p^2 - 11p + 4)}{(2p-2)T - (6p^2 - 10p + 4)} & 0 & \dots & 0 & -\frac{1}{2} & \frac{(p-2)T - (3p^2 - 6p + 2)}{(2p-2)T - (6p^2 - 10p + 4)} \end{bmatrix}'.$$

Nótese que, en este caso, la persistencia,  $\sum_{i=1}^p \alpha_i$ , tiende a  $\frac{(p-2)T - (4p^2 - 6p + 2)}{(p-1)T - (p^2 - 5p + 2)}$

Como resumen, se puede concluir que los efectos que las observaciones atípicas tienen en los estimadores MCO de los modelos ARCH( $p$ ) van a depender crucialmente de su magnitud, número y posición. El estimador del parámetro de escala,  $\omega$ , va a tender siempre a infinito con el tamaño de los atípicos. Sin embargo, los sesgos en los estimadores de  $\alpha_i$ ,  $i = 1, \dots, p$ , que miden la dependencia dinámica de la varianza, van a ser diferentes dependiendo del número y posición de los atípicos. Un único atípico sesga hacia cero los estimadores  $\hat{\alpha}_i \quad \forall i$ . Sin embargo, cuando hay  $k$  atípicos consecutivos, sus efectos dependerán de la relación entre  $k$  y  $p$ , el orden del modelo ARCH. Si el número de atípicos es suficientemente grande y mayor que  $p$ , el estimador del parámetro  $\alpha_1$  estará sesgado hacia uno y todos los demás  $\hat{\alpha}_i$  estarán sesgados hacia cero. No es posible establecer un resultado general para el caso  $k < p$ , dado que los sesgos son diferentes dependiendo de los valores concretos de  $k$  y  $p$ .

### 4.4.3 Sesgos en los estimadores MV

Obtener los sesgos causados por la presencia de observaciones atípicas en los estimadores MV de los modelos ARCH es complicado ya que dichos estimadores no tienen una expresión cerrada. Esto ha llevado a que todos los artículos que tratan este problema en la literatura lo hayan estudiado mediante experimentos de Monte Carlo, véase, por ejemplo, Mendes (2000) y Gregory y Reeves (2001).

Como ejemplo, considérese un modelo ARCH(1). Como se veía en la sección 4.4.1, los estimadores MV de los parámetros  $\omega$  y  $\alpha_1$  serán la solución del problema de optimización

$$\text{Minimizar}_{\omega, \alpha_1} \sum_{t=2}^T \left( \log(\omega + \alpha_1 y_{t-1}^2) + \frac{y_t^2}{(\omega + \alpha_1 y_{t-1}^2)} \right)$$

Si la serie está contaminada con un atípico de tamaño  $w_A$  en la observación  $\tau$ , la estimación consistirá en minimizar la siguiente expresión

$$\sum_{t \neq \tau} \left( \log(\omega + \alpha_1 y_{t-1}^2) + \frac{y_t^2}{(\omega + \alpha_1 y_{t-1}^2)} \right) + \log(\omega + \alpha_1 (y_\tau + w_A)^2) + \frac{(y_\tau + w_A)^2}{(\omega + \alpha_1 y_{\tau-1}^2)} + \frac{y_{\tau+1}^2}{(\omega + \alpha_1 (y_\tau + w_A)^2)}$$

Si  $w_A$  es grande y  $T$  moderado, la verosimilitud vendrá dominada por

$$\log(\omega + \alpha_1 (y_\tau + w_A)^2) + \frac{(y_\tau + w_A)^2}{(\omega + \alpha_1 y_{\tau-1}^2)} \quad (4.23)$$

En estos términos,  $\omega$  y  $\alpha_1$  deben neutralizar el efecto de  $w_A$ . Por tanto, para que el mínimo se alcance, el estimador de  $\alpha_1$  debe tender a cero para que  $\log(\hat{\omega} + \hat{\alpha}_1 (y_\tau + w_A)^2)$  sea lo menor posible y el estimador de  $\omega$  debe tender a infinito para hacer que el término  $\frac{(y_\tau + w_A)^2}{(\hat{\omega} + \hat{\alpha}_1 y_{\tau-1}^2)}$  sea también lo menor posible. Esto coincide con los sesgos encontrados para los estimadores MCO. No obstante, encontrar el óptimo de este problema de minimización no es sencillo debido a que la función objetivo es muy plana. Como ejemplo, en la figura 4.6, se ha representado la función objetivo para una serie, de tamaño  $T = 500$ , generada por un modelo ARCH(1), con parámetros  $\omega = 0.8$  y  $\alpha_1 = 0.2$  y para la misma serie contaminada con un único atípico de tamaños  $w_A = 5, 10, 25, 50$  y  $100$ . Como se puede observar, a medida que el tamaño del atípico crece, el valor de  $\omega$  que minimiza la función crece también, lo que se corresponde con lo que se esperaba viendo la expresión (4.23). En cuanto al valor de  $\alpha_1$  que minimiza la función, éste tiende a cero a medida que el tamaño del atípico aumenta. No obstante, nótese que la función a minimizar es muy plana, lo que hace difícil, para los métodos de optimización, encontrar la solución. Este hecho podría explicar los resultados que se muestran en la figura 4.7, donde se han representado los percentiles 25, 50 y 75 de los estimadores MV, basados en 500

réplicas, para el modelo ARCH(1) anterior, contaminado en primer lugar, por un único atípico, y luego por tres atípicos consecutivos. Como se puede observar, al igual que ocurre con los estimadores MCO, el estimador del parámetro  $\omega$  crece con el tamaño del atípico. En cuanto al estimador del parámetro  $\alpha_1$ , éste tiende a cero cuando la serie está contaminada con un único atípico, pero cuando hay tres atípicos consecutivos, la variabilidad es enorme para valores grandes de  $w_A$ . El estimador de  $\alpha_1$  podría tomar cualquier valor menor o mayor que el verdadero parámetro, lo cual podría ser debido a la dificultad de los métodos de optimización para minimizar la función objetivo, como ilustra la figura 4.6. Nótese, por otra parte, que para un tamaño del atípico fijo, los sesgos disminuyen con el tamaño muestral. Por ejemplo, para un único atípico de tamaño 7 desviaciones típicas, la mediana de los estimadores de  $\omega = 0.8$  y  $\alpha_1 = 0.2$  es 0.920 y 0.169 respectivamente cuando  $T = 500$  y 0.810 y 0.198 cuando  $T = 5000$ . Por tanto, para valores realistas de  $w_A$ , que son entre 5 y 10 desviaciones típicas, y tamaños muestrales grandes, los sesgos en los estimadores MV son menos importantes que en los estimadores MCO. Esto se puede apreciar también comparando las figuras 4.3 y 4.4 con 4.7. Nótese también un resultado interesante que sugiere la figura 4.7, y es que los errores de estimación parecen aumentar para tamaños del atípico pequeños mientras que a partir de un tamaño grande de  $w_A$ , el error estándar se hace cero, ya que como se aprecia en el gráfico, los percentiles 25, 50 y 75 de los estimadores coinciden. En consecuencia, parece que atípicos pequeños causan sesgos pequeños en la estimación puntual del parámetro pero causan sesgos importantes en el error estándar. Este resultado merece más atención y se estudiará con más detalle en el futuro.

#### Sesgos en los estimadores MV de los modelos GARCH(1,1)

La figura 4.8 contiene los percentiles 25, 50 y 75 de los estimadores MV, basados en 500 réplicas, de un modelo GARCH(1,1) con parámetros  $\omega = 0.1$ ,  $\alpha = 0.1$  y  $\beta = 0.8$ , contaminado en primer lugar con un único atípico y a continuación con tres atípicos

consecutivos. Como puede observarse en los gráficos, un único atípico parece sesgar hacia cero al estimador de  $\alpha$  y hacia uno al estimador de  $\beta$  mientras que el estimador de  $\omega$  parece no sentirse muy afectado por el atípico. En cuanto a tres atípicos consecutivos, parece que éstos sesgan a los estimadores de  $\omega$  y  $\alpha$  hacia un valor mayor que el verdadero, y al estimador de  $\beta$  hacia un valor menor. Parece, por tanto, que los sesgos causados por la presencia de atípicos en los estimadores MV de un modelo GARCH(1,1) dependen del número de atípicos, de su tamaño y, muy importante, de si los atípicos son aislados o consecutivos. Esta puede ser una razón para los resultados contradictorios encontrados en la literatura.

A partir de la representación ARCH( $\infty$ ) de los modelos GARCH(1,1), podría encontrarse una explicación a los sesgos causados por un único atípico. Considérese una serie  $y_t$  generada por un modelo GARCH(1,1) donde  $\sigma_0^2 = \frac{\omega}{1-\alpha-\beta}$ , es decir, el valor inicial para  $\sigma_t^2$  es la varianza marginal. Para un tamaño de muestra fijo,  $T$ , dicha serie puede aproximarse por el siguiente modelo ARCH( $T-1$ ):

$$\begin{cases} y_t = \varepsilon_t \sigma_t \\ \sigma_t^2 = \tilde{\omega} + \sum_{i=1}^{T-1} \tilde{\alpha}_i y_{t-i}^2 \end{cases}$$

donde  $\tilde{\omega} = \omega \left[ \frac{1+\beta^{T-1}}{1-\beta} + \frac{\beta^{T-1}}{1-\alpha-\beta} \right]$  y  $\tilde{\alpha}_i = \alpha\beta^{i-1}$  for  $i = 1, 2, \dots, T-1$ .

Si la serie observada está contaminada con un único atípico, a partir de la Proposición 2 se tendría que  $\lim_{w_A \rightarrow \infty} \hat{\tilde{\omega}} = \infty$  y  $\lim_{w_A \rightarrow \infty} \hat{\tilde{\alpha}}_i = \hat{\alpha}\hat{\beta}^{i-1} = 1/(T-2)$ ,  $\forall i$ . En particular, para  $i = 1$  se tendría que  $\hat{\alpha}$  tiende a  $1/(T-2)$  y por tanto,  $\hat{\beta}$  tiende a uno a medida que el tamaño del atípico tiende a infinito. En el límite,  $\hat{\alpha}$  está muy cerca de cero y  $\hat{\beta}$  está muy cerca de uno, lo que implica que la varianza marginal es infinito y  $\hat{\omega}$  no estaría identificado. Esto podría explicar que  $\hat{\omega}$  no se vea muy afectado por el atípico.

## 4.5 Modelización de la heterocedasticidad condicional en presencia de observaciones atípicas

Como se ha visto anteriormente, la presencia de observaciones atípicas puede distorsionar, notablemente, la inferencia y estimación en series temporales condicionalmente heterocedásticas. Por tanto, es importante corregir el efecto de los atípicos antes de analizar la dependencia dinámica de la varianza. Teniendo esto en cuenta, en esta sección se propone un procedimiento, en tres pasos, para modelizar la presencia simultánea de observaciones atípicas y heterocedasticidad condicional. El procedimiento propuesto es como sigue:

**Paso 1** Detectar y corregir las observaciones mayores que 7 desviaciones típicas.

**Paso 2** Estimar el modelo para la varianza condicional y calcular las observaciones estandarizadas en las que se detectan y corrigen los atípicos siguiendo los métodos tradicionales, puesto que éstas deberían ser homocedásticas.

**Paso 3** Reestimar el modelo de heterocedasticidad condicional.

En el Paso 1, la elección de 7 desviaciones típicas, como valor crítico, se ha hecho siguiendo dos argumentos. En primer lugar, por la desigualdad de Tchebychev se tiene que, para cualquier distribución de  $y_t$ :

$$P(|y_t - \bar{y}| \leq k\hat{\sigma}_y) \geq 1 - \frac{1}{k^2}$$

o lo que es lo mismo, en cualquier distribución, entre la media y  $k$  desviaciones típicas se encuentran, al menos, el  $100(1 - \frac{1}{k^2})$  por ciento de las observaciones. Por tanto, el 97.96% de las observaciones generadas a partir de cualquier distribución de media cero, son menores, en valor absoluto que 7 desviaciones típicas. El segundo argumento para elegir 7 desviaciones típicas, son los resultados que se muestran en la tabla 4.3, la cual

contiene los percentiles 95 de la distribución empírica del  $\max_{t=1,\dots,T} |y_t|$  para distintos tamaños muestrales, calculados a partir de 10000 réplicas de series  $y_t$  generadas por un modelo GARCH(1,1) con distintos valores de los parámetros. Como se puede observar en la tabla 4.3, para los valores de los parámetros que usualmente se estiman en series financieras reales, que son  $\alpha = 0.03, 0.05, 0.1, 0.15$ , y  $\beta = 0.8, 0.85, 0.9, 0.95$ , la probabilidad de obtener un valor mayor, en valor absoluto, que 7 desviaciones típicas es muy pequeña, pues el máximo valor generado por un modelo GARCH(1,1) es menor que 7 desviaciones típicas con una probabilidad muy alta. Este primer paso se hace para evitar sesgos importantes en la estimación de los parámetros. Como se ha visto en la sección anterior y se ilustrará luego con algunos ejemplos, los sesgos empiezan a ser importantes para atípicos mayores que 7 desviaciones típicas.

Una vez corregida la serie de los atípicos más extremos<sup>6</sup>, el segundo paso del procedimiento propuesto es estimar el modelo para las varianzas condicionales. Es importante notar que, en esta primera estimación, podrían estar presentes todavía observaciones atípicas que estarían sesgando la estimación. Por ello, lo siguiente es estandarizar las observaciones dividiendo por la varianza condicional estimada y detectar atípicos en dichas observaciones estandarizadas, que ya están próximas a la homocedasticidad. Una vez detectadas y corregidas las observaciones atípicas en las observaciones estandarizadas, el último paso es reestimar el modelo para la varianza condicional. Los Pasos 2 y 3 son iterativos, es decir, del Paso 3 se va de nuevo al Paso 2 hasta que ninguna observación sea detectada como atípico.

En relación al Paso 1 del procedimiento propuesto, es importante notar que, dependiendo del valor crítico elegido, se corre el riesgo de corregir observaciones que no son atípicos sino observaciones correspondientes a periodos donde la varianza condicional es mayor que la marginal, y por ello 7 desviaciones típicas parece bastante razonable.

---

<sup>6</sup>El modo tradicional de corregir la serie es substituir la observación atípica por la media condicional de la serie.



Para ilustrar el efecto del valor crítico, considérense los resultados obtenidos por Verhoeven y McAleer (2000), quienes estudian los efectos de observaciones atípicas en el modelo AR(1)-GARCH(1, 1), definiendo atípicos como aquellas observaciones mayores, en valor absoluto, que 3 desviaciones típicas, sin tener en cuenta la heterocedasticidad condicional. Analizando rendimientos diarios de cinco series financieras, concluyen que cuando se eliminan las observaciones atípicas, el estadístico de Ljung-Box disminuye significativamente y que los atípicos causan un incremento significativo en el parámetro estimado  $\hat{\alpha}$  del modelo GARCH(1, 1), y un decaimiento del parámetro estimado  $\hat{\beta}$ . La figura 4.9 ilustra los efectos que puede tener corregir por observaciones mayores que 3.5 desviaciones típicas, si no se tiene en cuenta que además hay efectos ARCH. En dicho gráfico se han representado los resultados de un estudio de Monte Carlo en donde se han simulado 1000 réplicas de series generadas por un modelo GARCH(1,1) con parámetros  $\omega = 0.1$ ,  $\alpha = 0.1$  y  $\beta = 0.8$ , para tamaños muestrales  $T = 500, 1000$  y  $5000$ . La primera fila del gráfico contiene los resultados para  $T = 500$ , la segunda y la tercera fila contienen los mismos para  $T = 1000$  y  $T = 5000$  respectivamente. En el eje de abscisas se han representado los parámetros estimados obtenidos a partir de la serie original simulada, el eje de ordenadas contiene las estimaciones obtenidas a partir de la serie corregida por observaciones mayores que 3.5 desviaciones típicas. Como se puede observar, hay puntos situados sobre la diagonal, que significa que las estimaciones son las mismas, lo que sucede, por ejemplo, cuando ninguna observación de la serie simulada es mayor que  $3.5\sigma$ . En los casos en que los parámetros estimados son diferentes, puede observarse como para la mayoría de las series corregidas,  $\omega$  y  $\alpha$  se estiman menores y  $\beta$  se estima mayor que en las series originales. Esto podría ser lo que sucede en Verhoeven y McAleer (2000). Sin embargo, otra posibilidad para los resultados encontrados por estos autores es que existan, en los datos, atípicos consecutivos, los cuales podrían estar exagerando la heterocedasticidad y, por tanto, al corregirlos,  $\hat{\alpha}$  disminuiría y  $\hat{\beta}$  aumentaría.

Una alternativa para modelizar la heterocedasticidad condicional cuando hay datos atípicos es mediante métodos robustos. Algunos artículos que proponen estimadores robustos para los modelos ARCH son Sakata y White (1998) y Muler y Yohai (2002). No obstante, estos métodos suelen ser muy pesados computacionalmente y menos eficientes que los métodos tradicionales no robustos. Otra ventaja del procedimiento propuesto en esta sección, además de su sencillez en la implementación, es que permite identificar las observaciones atípicas aportando más información sobre la serie de interés.

## 4.6 Ejemplos

Para ilustrar los resultados obtenidos anteriormente, en esta sección se analizan varias series reales. En primer lugar, se consideran los rendimientos del índice S&P 500, observado diariamente desde el 4 de Enero de 1982 hasta el 23 de Septiembre de 1994 (3219 observaciones) y el tipo de cambio Dólar estadounidense frente al Yen japonés, observado también diariamente desde el 2 de Enero de 1973 hasta el 29 de Enero de 2002 (7320 observaciones)<sup>7</sup>. En las figuras 4.10 y 4.11 se han representado gráficamente las series de rendimientos y el correlograma de los mismos al cuadrado. Ambas series se han corregido por las observaciones más extremas para ilustrar el efecto que dichas observaciones causan en el correlograma de los rendimientos al cuadrado. En primer lugar se han detectado las observaciones mayores, en valor absoluto, que 10 desviaciones típicas<sup>8</sup> y se han corregido sustituyéndolas por su media condicional, que en ambos casos coincide con la media marginal ya que ambas series son incorreladas. Para la serie del S&P 500 hay solamente una observación mayor que 10 desviaciones

---

<sup>7</sup>La serie del S&P 500 es la analizada en Pérez y Ruiz (2002) y los datos para el tipo de cambio dólar/yen han sido obtenidos de la página web <http://pacific.commerce.ubc.ca/xr/> por cortesía del Prof. Werner Antweiler, University of British Columbia, Vancouver, Canada.

<sup>8</sup>El valor de  $10\sigma$  se ha elegido de modo arbitrario para ilustrar el efecto de los atípicos más extremos.

típicas<sup>9</sup>, que es la correspondiente al 19 de Octubre de 1987, conocido también como “October black monday”, cuando tuvo lugar la caída más fuerte de la historia de la bolsa de Wall Street. Como se puede observar en la figura 4.10, el hecho de corregir la serie por una única observación hace que la estructura en los cuadrados aparezca más claramente. Este caso es un ejemplo en donde una observación atípica está, posiblemente, enmascarando la heterocedasticidad condicional presente en la serie. Si se corrige la serie por las observaciones mayores que 7 desviaciones típicas, el correlograma se ve afectado de nuevo, apareciendo significativas correlaciones de orden mayor. Las diferencias al corregir la serie por observaciones mayores que 5 desviaciones típicas no son tan dramáticas como en los casos anteriores y apenas hay diferencias entre el correlograma de los rendimientos al cuadrado de la serie corregida por observaciones mayores que 7 desviaciones típicas y el de la serie corregida por observaciones mayores que 5 desviaciones típicas.

De un modo similar pueden observarse los efectos que las observaciones atípicas pueden tener en la identificación de la heterocedasticidad condicional cuando se analiza la serie de rendimientos del tipo de cambio dólar/yen representada en la figura 4.11. Al igual que en el ejemplo anterior, en primer lugar se han detectado las observaciones mayores, en valor absoluto, que 10 desviaciones típicas. A continuación se han corregido las observaciones mayores que 7 desviaciones típicas y por último las mayores que 5 desviaciones típicas. Los efectos que una única observación puede tener en el correlograma se ponen de manifiesto de nuevo en este gráfico. De nuevo, no parece haber grandes diferencias entre el correlograma de los rendimientos al cuadrado de la serie corregida por observaciones mayores que 7 desviaciones típicas y el de la serie corregida por observaciones mayores que 5 desviaciones típicas.

Por lo tanto, como ilustran estos ejemplos, en las series reales puede haber observaciones extremas, que pueden considerarse atípicos, que están causando sesgos en el

---

<sup>9</sup>Dicha observación es exactamente 22 veces la desviación típica.

correlograma de las observaciones al cuadrado y que pueden afectar de manera decisiva a la identificación de la heterocedasticidad condicional. No obstante, para estas dos series en particular, las observaciones atípicas no afectan al diagnóstico de los cuatro contrastes de homocedasticidad condicional, los cuales rechazan la hipótesis nula en todos los casos.

A continuación, para ilustrar el procedimiento propuesto en la sección anterior, se ha aplicado dicho procedimiento a las 12 series analizadas en el Capítulo 2. En primer lugar se han aplicado a dichas series los cuatro contrastes de heterocedasticidad condicional descritos en la sección 4.3.1 obteniendo que para todas las series se rechaza la hipótesis nula de homocedasticidad condicional con los contrastes no robustos, es decir, el de Lagrange, el de McLeod y Li (1983) y el propuesto por Peña y Rodríguez (2002). Sin embargo, el contraste robusto propuesto por Van Dijk et al. (1999) rechaza la hipótesis nula solamente para siete de las doce series. No la rechaza para los tipos de cambio US-CAN, US-GER, US-JAP, US-SSF, ni para el índice bursátil de Bombay, lo que podría ser consecuencia de su baja potencia.

La tabla 4.4 contiene las observaciones detectadas como atípicas mediante el procedimiento propuesto. En negrita se encuentran los atípicos más grandes, detectados en el primer paso, es decir, aquellas observaciones mayores que 7 desviaciones típicas y que son las que se corrigen en primer lugar para evitar que causen sesgos importantes en la estimación de los parámetros. Como justificación adicional a la elección de 7 desviaciones típicas en el Paso 1 del procedimiento, se ha realizado el siguiente experimento: se han calculado los percentiles 95 de la distribución empírica del  $\max_{t=1,\dots,T} |y_t|$  para los tamaños muestrales correspondientes a las 12 series, calculados a partir de 10000 réplicas de series  $y_t$  generadas por un modelo GARCH(1,1) gaussiano con los valores de los parámetros estimados para las 12 series que aparecen en la tabla 2.2. Los percentiles correspondientes a las 12 series son 8.34, 5.77, 5.61, 6.41, 5.01, 5.49, 5.68, 9.33, 5.58, 9.17, 8.56 y 7.03 respectivamente. Por tanto, parece que 7 desviaciones típicas es un

valor razonable, aunque para algunas series el percentil 95 del  $\max_{t=1,\dots,T} |y_t|$  es hasta de 9 desviaciones típicas.

La tabla 4.5 contiene los estadísticos descriptivos de las series corregidas. Comparando ésta con la tabla 2.1 se puede observar cómo en las series corregidas, la desviación típica disminuye, al igual que la curtosis. En cuanto al coeficiente de asimetría, nótese que para las series del tipo de cambio US-SWK y los índices del Dow Jones y S&P 500 la asimetría negativa que aparece en las series originales está completamente explicada por las observaciones atípicas, ya que una vez que éstas son tenidas en cuenta, el coeficiente de asimetría no es significativo. En cuanto al coeficiente de correlación de orden uno de las observaciones al cuadrado, no parece haber un comportamiento uniforme a todas las series, para algunas aumenta al tener en cuenta los atípicos; véase, por ejemplo, los índices de Amsterdam y Bombay, mientras que para otras disminuye, como por ejemplo, los índices Dow Jones y S&P 500. Esto es lo esperable según el tipo de atípicos. En la tabla 4.4 puede comprobarse que los atípicos corregidos en las dos primeras series son aislados, mientras que en las dos últimas se han corregido atípicos consecutivos. Finalmente, el estadístico  $Q_2(20)$  aumenta significativamente cuando se tienen en cuenta las observaciones atípicas.

La tabla 4.6 contiene las estimaciones de los modelos GARCH(1,1) gaussianos para las series corregidas por los atípicos detectados mediante el procedimiento descrito y que se detallan en la tabla 4.4. Comparando estos estimadores con los de la tabla 2.2 se concluye que, los sesgos en los parámetros dependen del tamaño de los atípicos y de si éstos son aislados o consecutivos. Por ejemplo, en la serie del índice de Bombay, parece claro que la presencia de atípicos aislados está sesgando el parámetro  $\alpha$  hacia cero y  $\beta$  hacia uno. Nótese además que la verosimilitud es mayor para las series corregidas que para las originales. Es importante destacar también que, a pesar de que numéricamente los estimadores de los parámetros son diferentes en las series corregidas y en las originales, estadísticamente no lo son en algunos casos, ya que al 95% de confianza

no se puede rechazar que los valores estimados en las tablas 2.2 y 4.6 sean diferentes. Para el índice de Bombay, los estimadores de  $\alpha$  y  $\beta$  son significativamente diferentes cuando se estiman en las series originales y corregidas. En cuanto al parámetro  $\omega$ , hay diferencias significativas para las series de los tipos de cambio US-ESP, US-JAP y US-SSF.

## 4.7 Conclusiones

En este capítulo se ha abordado el problema de la presencia simultánea de heterocedasticidad condicional y observaciones atípicas en series temporales. Se han estudiado los efectos causados por los atípicos en la identificación de la heterocedasticidad condicional encontrando que, un único atípico sesga hacia cero los coeficientes de correlación de las observaciones al cuadrado, y rachas de atípicos generan correlaciones en los cuadrados de la serie, lo que puede llevar a los contrastes de homocedasticidad condicional a rechazar la hipótesis nula. Se han obtenido los sesgos asintóticos causados por los atípicos en los estimadores MCO de procesos ARCH( $p$ ), concluyendo que un único atípico sesga hacia cero los parámetros que miden la heterocedasticidad y atípicos consecutivos pueden sobreestimarlos, y los sesgos dependen, en este caso, del número de atípicos y del orden del proceso. Mediante estudios de simulación se han ilustrado tales efectos en muestras finitas, tanto en los estimadores MCO como en los estimadores MV de modelos ARCH( $p$ ) y GARCH(1,1). Finalmente, se ha propuesto un procedimiento, en tres pasos, para modelizar la heterocedasticidad condicional en presencia de observaciones atípicas y se ha ilustrado con las doce series descritas en el Capítulo 2.

4.8 Apéndice de tablas y figuras

Figura 4.1: Funciones  $\psi$  y  $w_y$

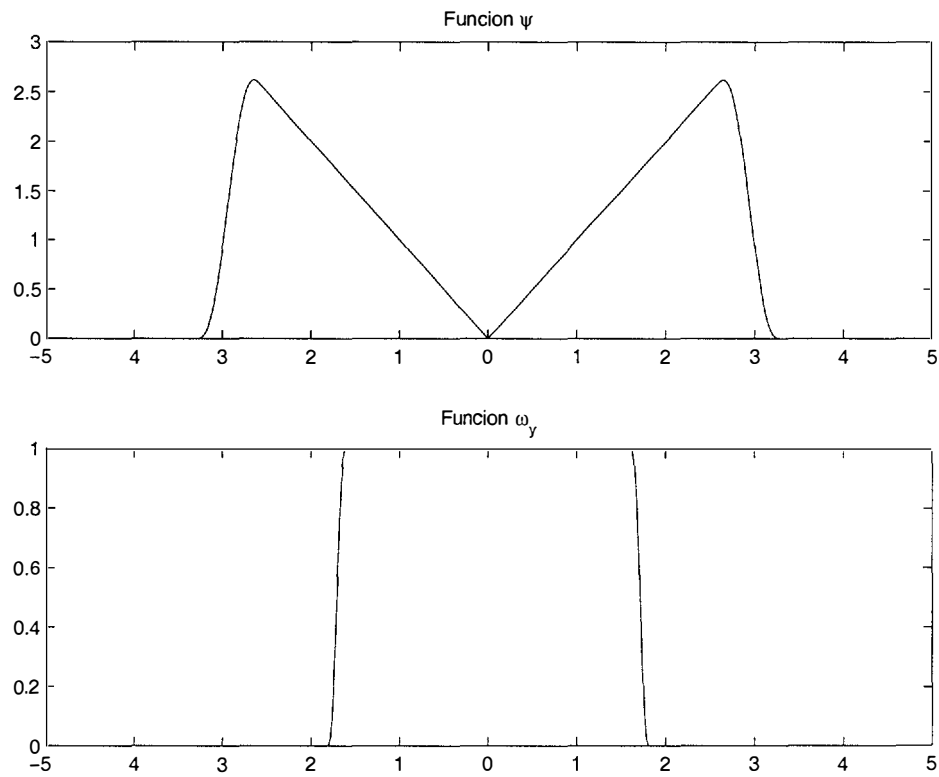


Tabla 4.1: Efectos causados por la presencia de observaciones atípicas en el diagnóstico de la heterocedasticidad condicional basados en 1000 réplicas para cuatro modelos con atípicos consecutivos (AC) y aislados (AI)

Serie	Estadísticos descriptivos					Contrastes de homocedasticidad (Porcentaje de rechazos)			
	Media	D.T.	Simetría	Curtosis	$r_2(1)$	LM	McLL	PR	Robusto
Normal ( $a_t$ )	-0.0002	0.9993	-0.0016	2.9936	-0.0006	4.00	5.20	5.50	5.90
AC ( $a_t^c$ )	0.0098	1.0716	0.1989*	5.7333*	0.3783*	99.70	98.90	99.70	6.00
AI ( $a_t^a$ )	0.0298	1.0709	0.5857*	5.6212*	-0.0043	1.90	0.60	0.70	7.00
GARCH ( $y_t$ )	-0.0017	0.9962	-0.0008	3.2375*	0.1180*	63.70	78.60	79.00	49.10
AC ( $y_t^c$ )	0.0083	1.0689	0.2017*	6.0154*	0.4103*	99.80	99.70	99.90	51.10
AI ( $y_t^a$ )	0.0283	1.0686	0.6072*	5.9654*	0.0425	16.40	13.20	13.90	50.30
EGARCH ( $x_t$ )	-0.0027	0.9977	0.0193	3.2195*	0.0557*	26.40	61.00	53.80	27.80
AC ( $x_t^c$ )	0.0073	1.0702	0.2233*	6.0632*	0.3820*	99.40	99.40	99.60	27.20
AI ( $x_t^a$ )	0.0273	1.0708	0.6385*	6.0502*	0.0191	6.20	11.30	7.60	28.40
ARSV ( $z_t$ )	-0.0017	0.9722	-0.0118	6.7273*	0.1960*	89.10	99.50	99.20	79.60
AC ( $z_t^c$ )	0.0083	1.0505	0.2729*	10.3783*	0.3986*	99.80	99.90	100	80.00
AI ( $z_t^a$ )	0.0283	1.0504	0.8241*	10.3550*	0.0995*	46.70	62.30	58.90	80.70

\* Significativo al 5%.



Figura 4.2: Correlogramas de los cuadrados de distintas series con y sin observaciones atípicas

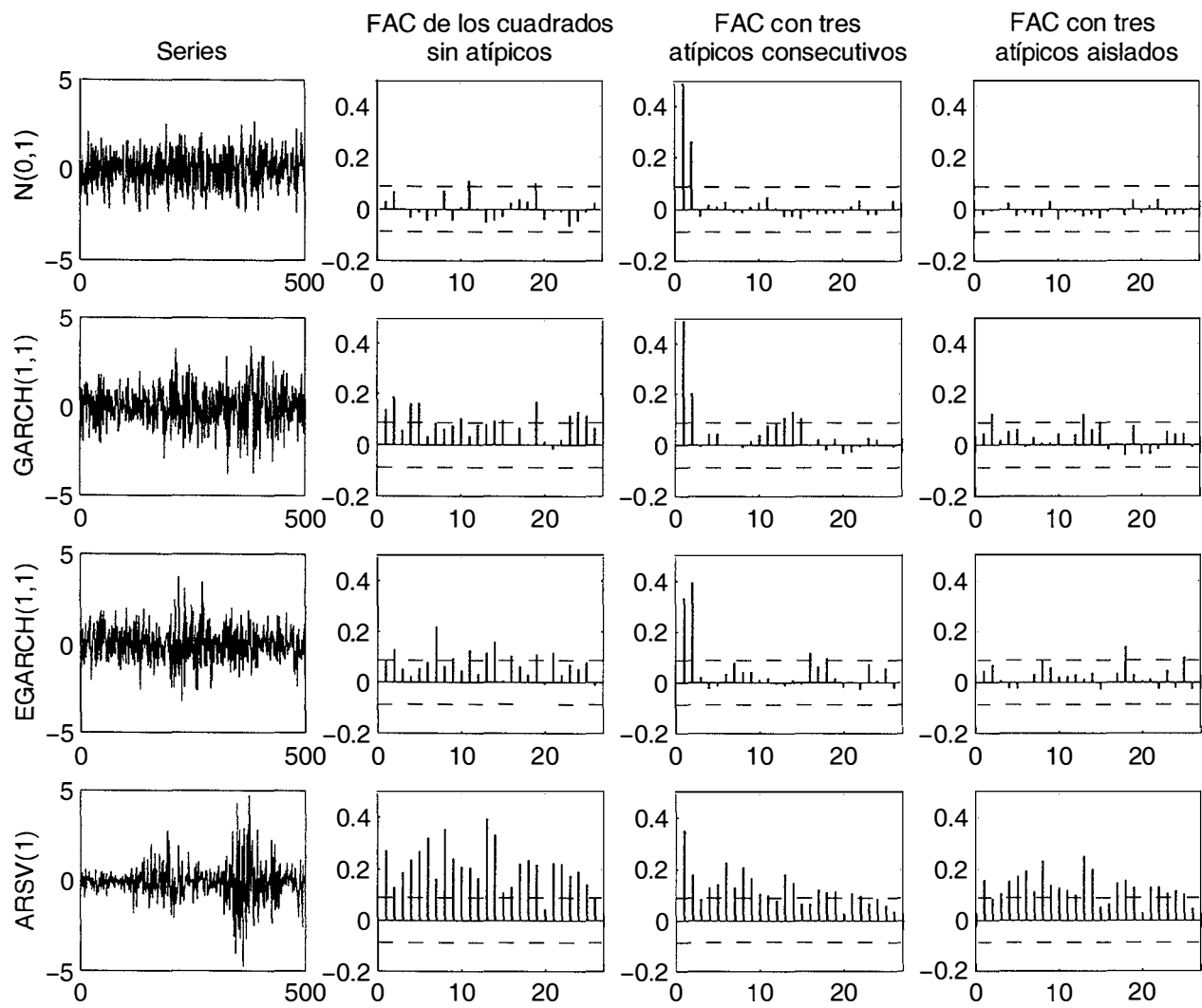


Figura 4.3: Cuantiles de los estimadores MCO de modelos ARCH( $p$ ) con un único atípico

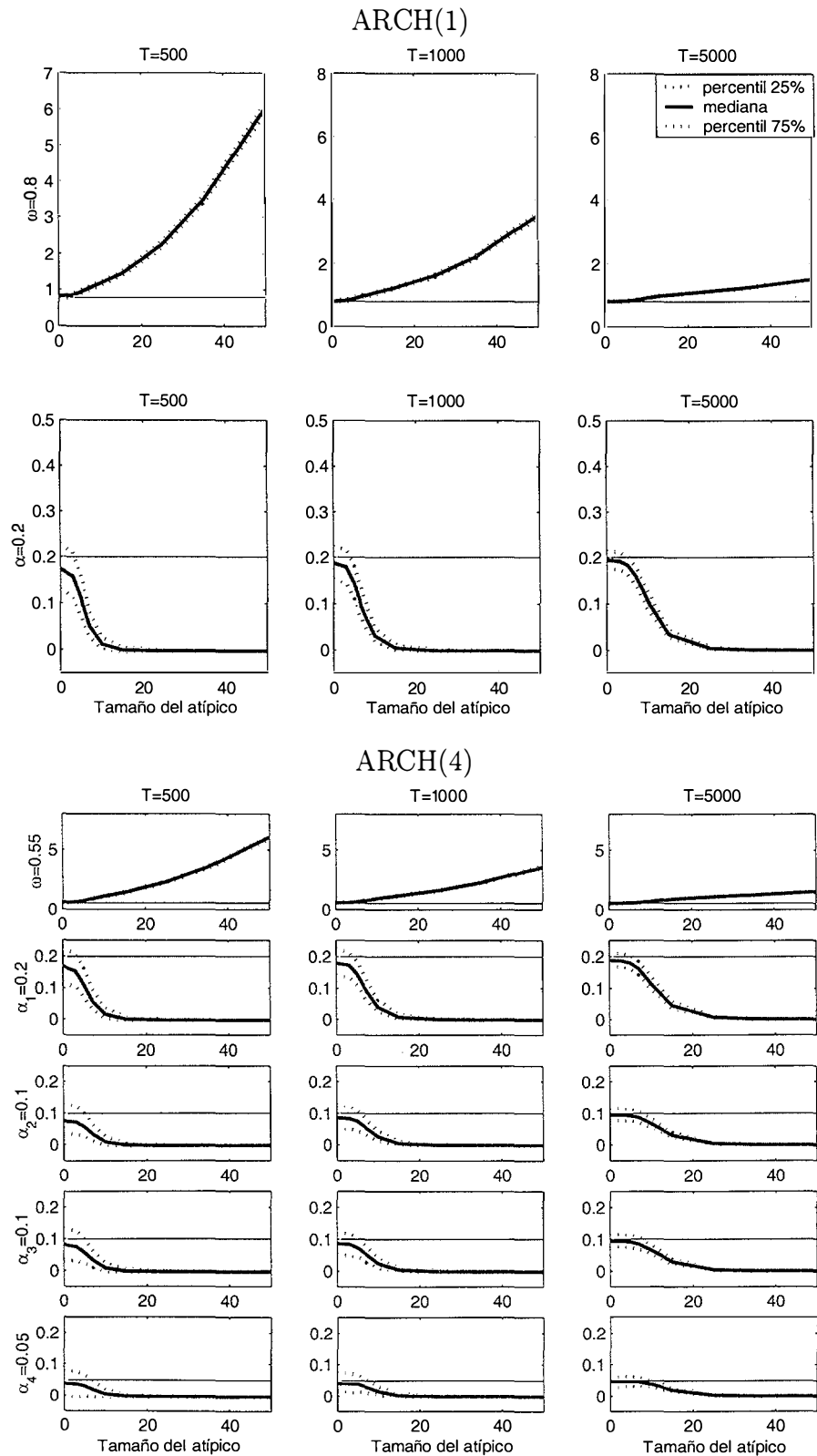


Figura 4.4: Cuantiles de los estimadores MCO de un modelo ARCH(1) con 2 y 20 atípicos consecutivos

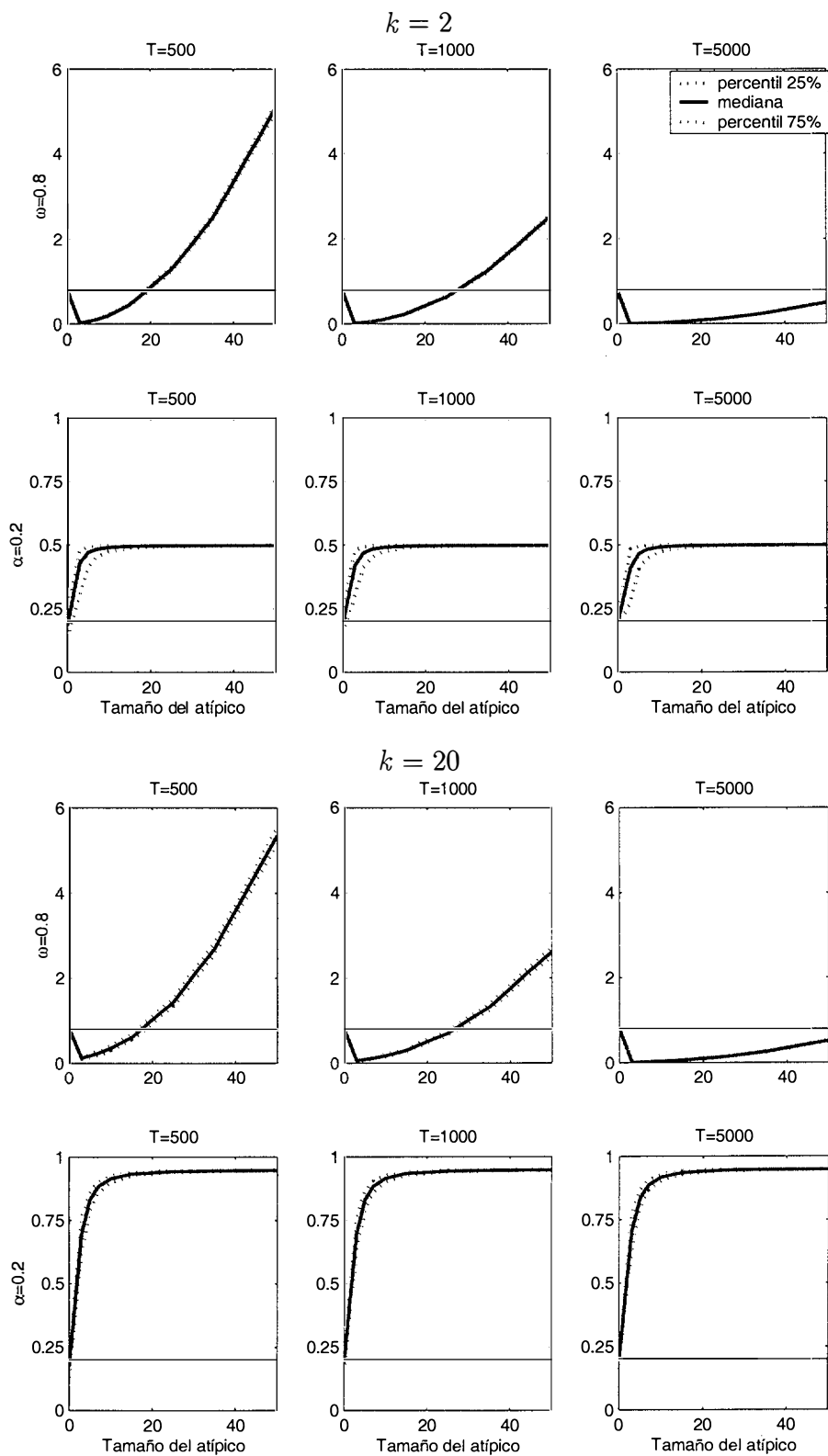


Figura 4.5: Cuantiles de los estimadores MCO de un modelo ARCH(4) con 20 y 2 atípicos consecutivos

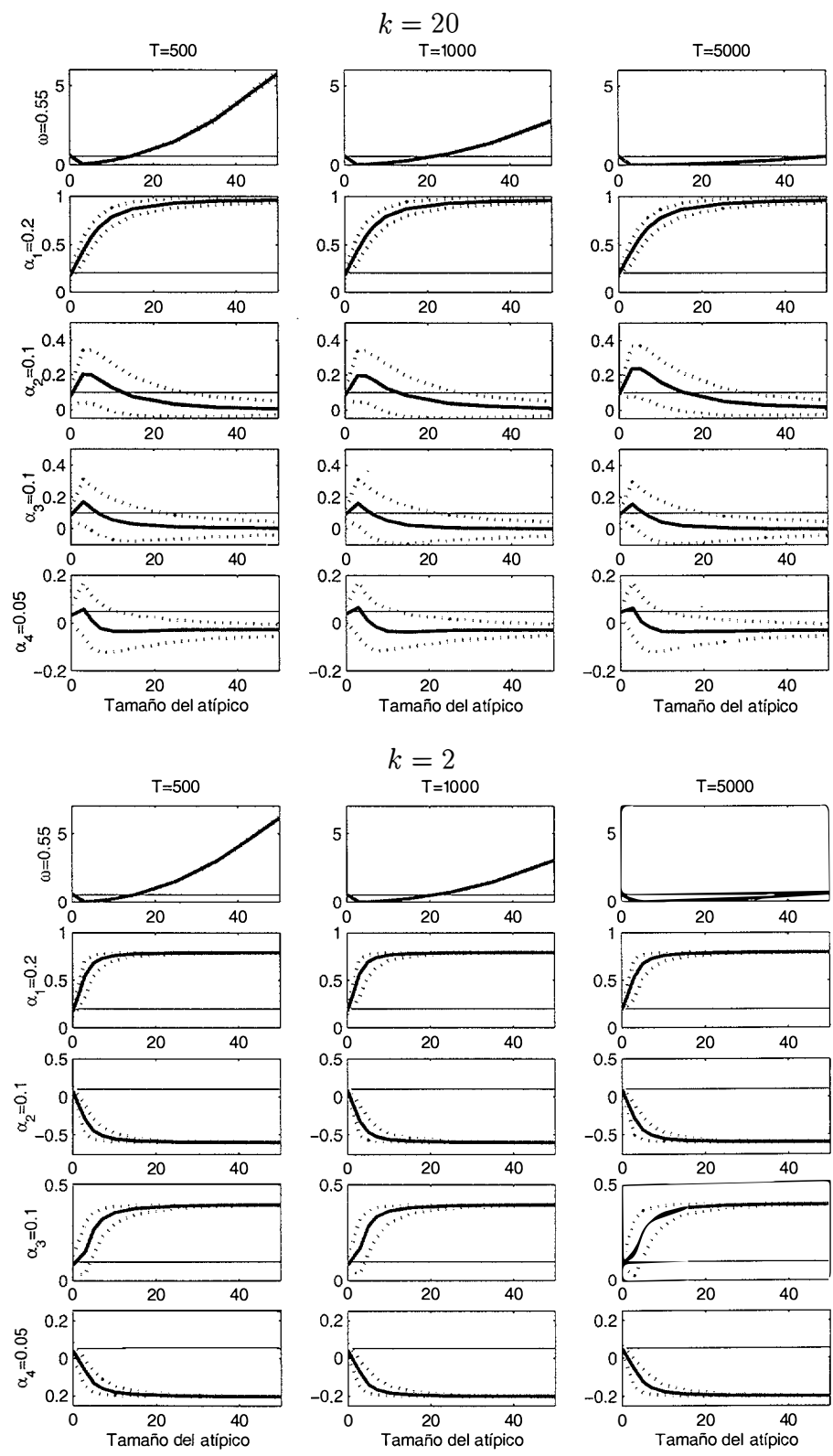


Tabla 4.2: Límites de los estimadores MCO de modelos ARCH( $p$ ) cuando el tamaño del atípico tiende a infinito

$k$	2	3	4
ARCH(3)	$\hat{\omega} \rightarrow \infty$ $\hat{\alpha}_1 \rightarrow \frac{3T-25}{4T-28}$ $\hat{\alpha}_2 \rightarrow -\frac{1}{2}$ $\hat{\alpha}_3 \rightarrow \frac{T-11}{4T-28}$	$\hat{\omega} \rightarrow \infty$ $\hat{\alpha}_1 \rightarrow \frac{3T-27}{4T-30}$ $\hat{\alpha}_2 \rightarrow 0$ $\hat{\alpha}_3 \rightarrow \frac{-T+3}{4T-30}$	$\hat{\omega} \rightarrow \infty$ $\hat{\alpha}_1 \rightarrow \frac{5T-47}{6T-50}$ $\hat{\alpha}_2 \rightarrow 0$ $\hat{\alpha}_3 \rightarrow \frac{T+3}{6T-50}$
ARCH(4)	$\hat{\omega} \rightarrow \infty$ $\hat{\alpha}_1 \rightarrow \frac{4T-40}{5T-44}$ $\hat{\alpha}_2 \rightarrow \frac{-3T+24}{5T-44}$ $\hat{\alpha}_3 \rightarrow \frac{2T-20}{5T-44}$ $\hat{\alpha}_4 \rightarrow \frac{-T+4}{5T-44}$	$\hat{\omega} \rightarrow \infty$ $\hat{\alpha}_1 \rightarrow \frac{5T-56}{6T-60}$ $\hat{\alpha}_2 \rightarrow 0$ $\hat{\alpha}_3 \rightarrow -\frac{1}{2}$ $\hat{\alpha}_4 \rightarrow \frac{2T-26}{6T-60}$	$\hat{\omega} \rightarrow \infty$ $\hat{\alpha}_1 \rightarrow \frac{4T-48}{5T-52}$ $\hat{\alpha}_2 \rightarrow 0$ $\hat{\alpha}_3 \rightarrow 0$ $\hat{\alpha}_4 \rightarrow \frac{T+4}{5T-52}$
ARCH(5)	$\hat{\omega} \rightarrow \infty$ $\hat{\alpha}_1 \rightarrow \frac{5T-61}{6T-66}$ $\hat{\alpha}_2 \rightarrow -\frac{2}{3}$ $\hat{\alpha}_3 \rightarrow \frac{3T-39}{6T-66}$ $\hat{\alpha}_4 \rightarrow -\frac{1}{3}$ $\hat{\alpha}_5 \rightarrow \frac{T-17}{6T-66}$	$\hat{\omega} \rightarrow \infty$ $\hat{\alpha}_1 \rightarrow \frac{7T-89}{8T-94}$ $\hat{\alpha}_2 \rightarrow \frac{-T+5}{4(4T-47)}$ $\hat{\alpha}_3 \rightarrow -\frac{1}{2}$ $\hat{\alpha}_4 \rightarrow \frac{7T-89}{4(4T-47)}$ $\hat{\alpha}_5 \rightarrow \frac{-T+5}{2(4T-47)}$	$\hat{\omega} \rightarrow \infty$ $\hat{\alpha}_1 \rightarrow \frac{7T-99}{8T-104}$ $\hat{\alpha}_2 \rightarrow 0$ $\hat{\alpha}_3 \rightarrow 0$ $\hat{\alpha}_4 \rightarrow -\frac{1}{2}$ $\hat{\alpha}_5 \rightarrow \frac{3T-47}{8T-104}$

Figura 4.6:  $-2 * \log$ -verosimilitud para una serie generada por un modelo ARCH(1) con parámetros  $\omega = 0.8$  y  $\alpha = 0.2$  y que se contamina con un atípico de tamaño  $w_A = 5, 10, 25, 50$  y  $100$ .

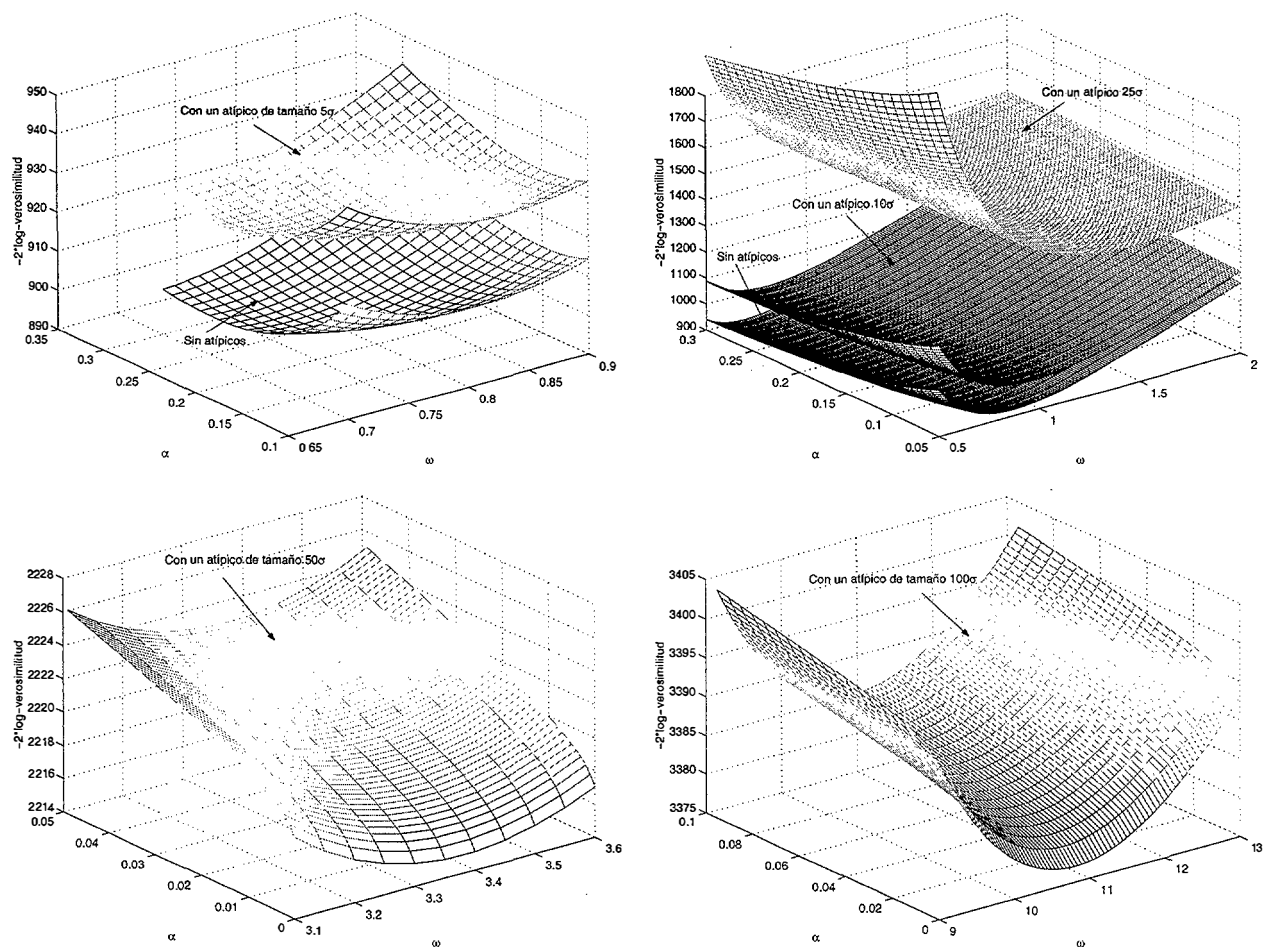


Figura 4.7: Cuantiles de los estimadores MV de un modelo ARCH(1) con atípicos

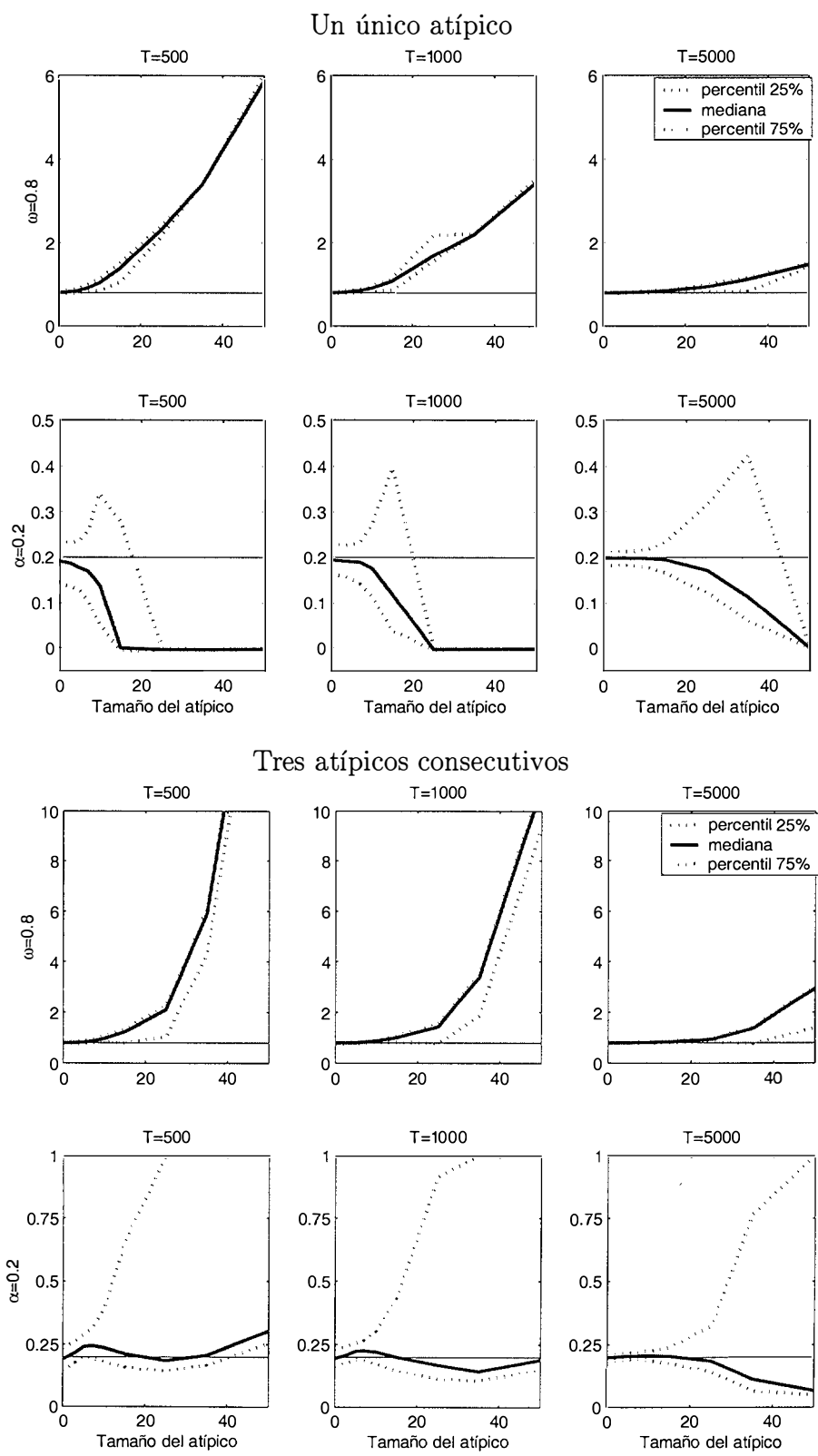


Figura 4.8: Cuantiles de los estimadores MV de un modelo GARCH(1,1) con atípicos

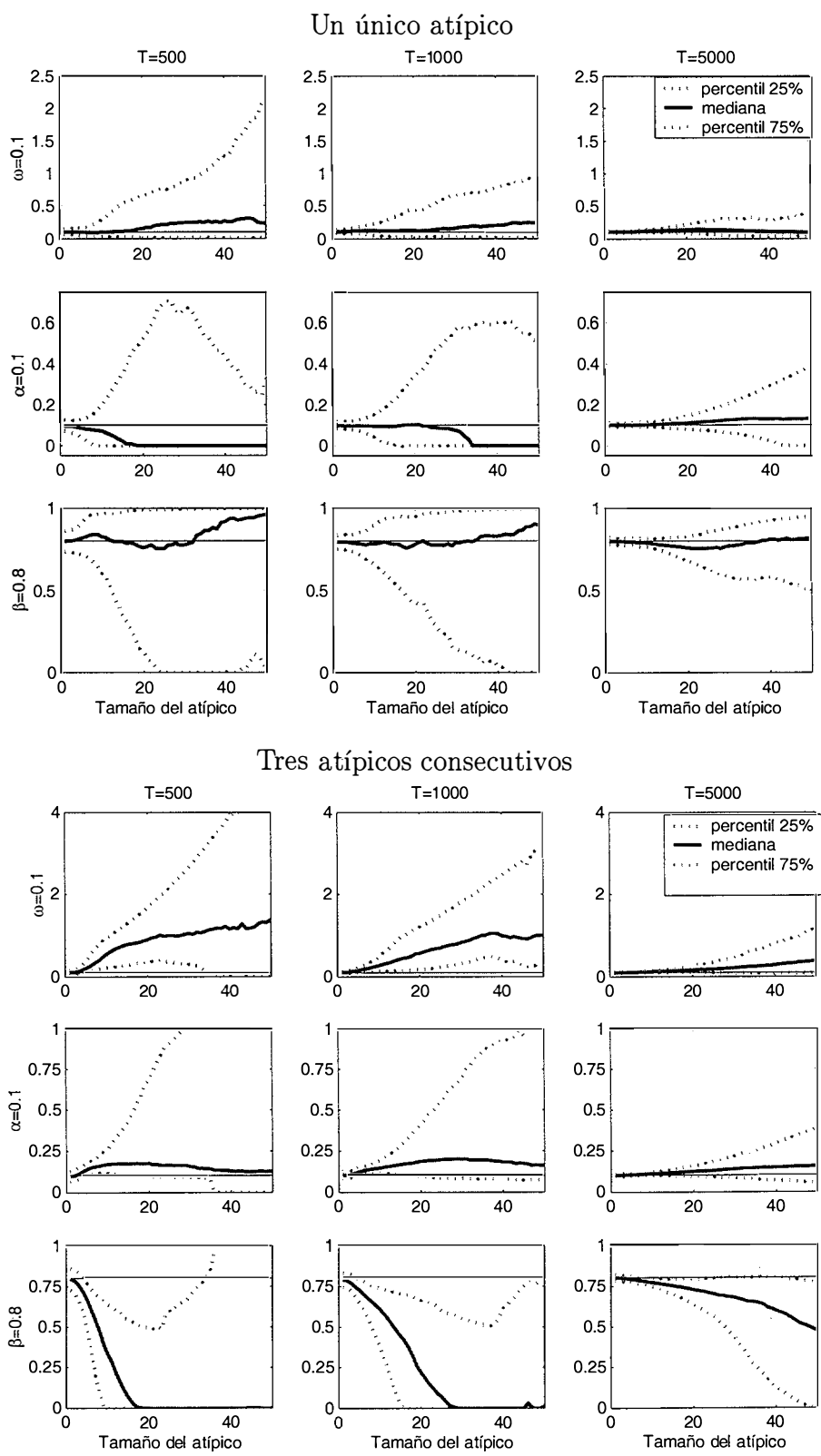




Tabla 4.3: Percentiles empíricos de  $\max |y_t|$  al 95% para  $y_t$  GARCH(1,1)

GARCH(1,1)			Tamaño muestral				
$\omega$	$\alpha$	$\beta$	$T = 500$	$T = 1000$	$T = 2500$	$T = 5000$	$T = 10000$
0.10	0.10	0.80	4.5386	4.8833	5.3970	5.7757	6.0926
0.05	0.10	0.85	5.0370	5.5263	6.2954	6.8355	7.3618
0.15	0.05	0.80	3.9812	4.1784	4.4355	4.6341	4.8096
0.10	0.05	0.85	4.0462	4.2401	4.5246	4.7215	4.9115
0.05	0.05	0.90	4.2113	4.4344	4.7662	5.0088	5.2373
0.10	0.02	0.88	3.8933	4.0613	4.2823	4.4668	4.6137
0.03	0.02	0.95	3.9558	4.1302	4.3675	4.5574	4.7288
0.05	0.20	0.75	7.3117	8.6032	10.9001	12.8283	14.7093
0.10	0.20	0.70	6.1848	6.9952	8.3290	9.4298	10.5697
0.05	0.15	0.80	6.1630	7.0184	8.4677	9.5240	10.6232
0.10	0.15	0.75	5.2885	5.8467	6.6888	7.3722	8.0497

Figura 4.9: Estimadores GARCH(1,1) basados en 1000 series corregidas por observaciones mayores, en valor absoluto, que  $3.5\sigma$

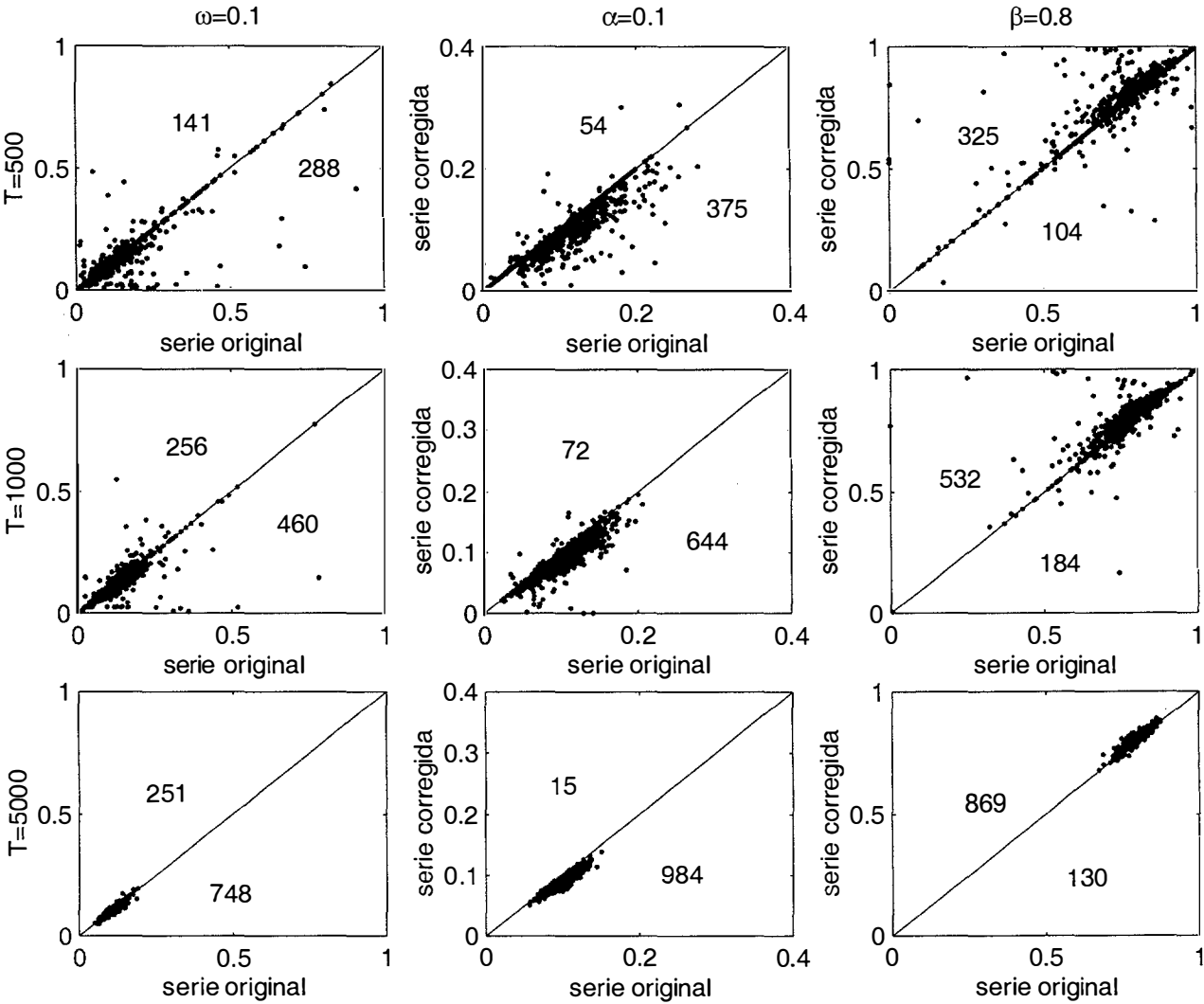


Figura 4.10: Correlogramas de los cuadrados de los rendimientos del índice S&P 500 con y sin observaciones atípicas

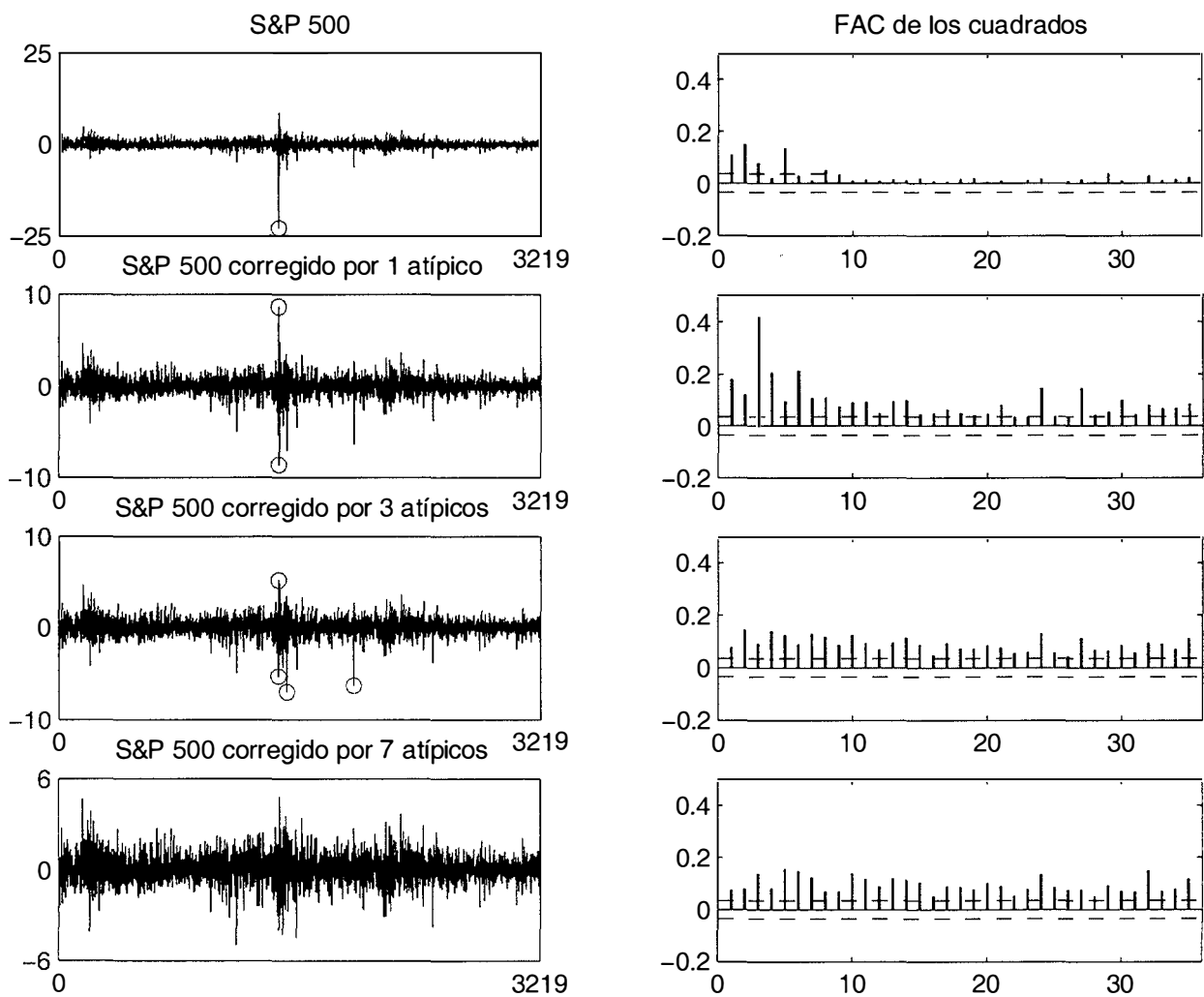


Figura 4.11: Correlogramas de los cuadrados de los rendimientos del tipo de cambio Dólar estadounidense y Yen Japonés con y sin observaciones atípicas

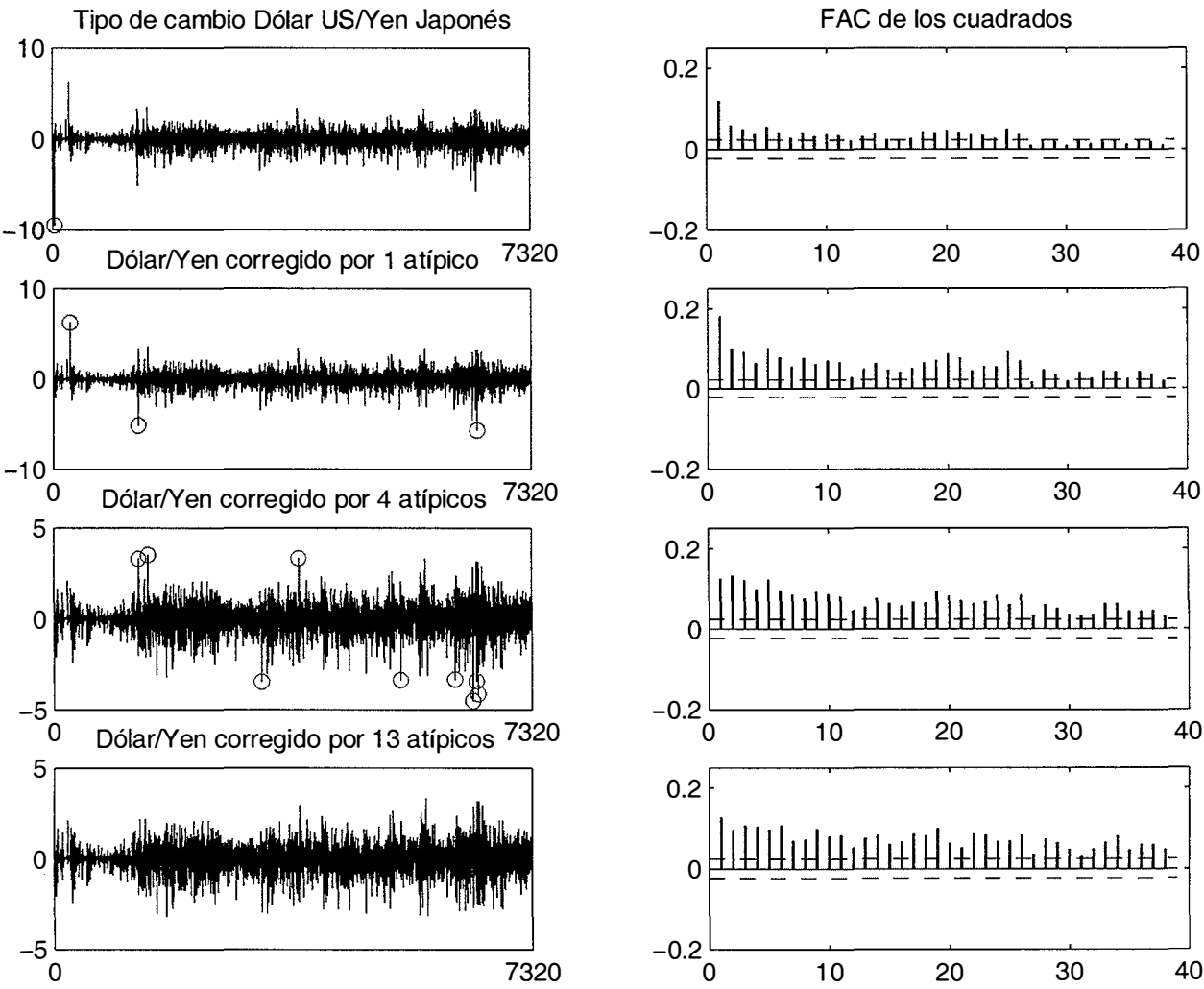


Tabla 4.4: Observaciones atípicas

[illegible]

Tabla 4.5: Estadísticos descriptivos de varias series diarias de rendimientos financieros corregidas por la media y por las observaciones atípicas

	US CAN	US ESP	US GER	US JAP	US SSF	US SWK	US UKP	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
T	1963	1963	1963	1963	1963	1963	1963	1261	1262	2728	1982	2888
Media	0.0078	0.0259	0.0186	0.0031	0.0127	0.0167	-0.0007	0.0127	0.0139	0.0079	-0.0011	0.0126
D.T.	0.3047	0.6061	0.6014	0.7570	0.6744	0.6174	0.5012	1.3209	1.7282	0.8873	1.2194	0.8532
Simetría	0.1346*	0.2263*	0.2192*	0.4101*	0.2251*	0.0844	0.1080	0.1949*	0.1172	0.0386	0.3098*	0.0025
Curtosis	5.5281*	4.2746*	4.5419*	7.0506*	4.5429*	4.0628*	5.1474*	5.5364*	5.2649*	4.9874*	6.3231*	5.3564*
$r(1)$	0.0040	0.0370	0.0100	0.0150	0.0080	0.0010	0.0050	0.0290	0.0100	0.0150	0.0050	0.0380
$Q(20)$	28.09	21.47	20.20	20.61	19.52	18.48	30.70	47.65*	32.45*	55.99*	39.77*	64.68*
Autocorrelaciones de $y_t^2$												
$r_2(1)$	0.1410*	0.1060*	0.1040*	0.2960*	0.1450*	0.0650*	0.1100*	0.2200*	0.2100*	0.1250*	0.2590*	0.1050*
$r_2(2)$	0.0860*	0.0740*	0.0290	0.1250*	0.0200	0.0680*	0.0980*	0.3080*	0.2400*	0.1350*	0.2410*	0.1300*
$r_2(5)$	0.1490*	0.0510*	0.0740*	0.1200*	0.0730*	0.0430	0.0820*	0.1590*	0.1720*	0.1200*	0.2260*	0.1420*
$r_2(10)$	0.0810*	0.0730*	0.1090*	0.0400	0.0680*	0.0570*	0.0810*	0.1510*	0.0930*	0.1040*	0.2810*	0.1610*
$Q_2(20)$	475*	169*	182*	509*	217*	162*	301*	882*	387*	721*	1842*	1268*

T: Tamaño muestral.

$r(\tau)$ : Autocorrelación de orden  $\tau$  de las observaciones  $y_t$ .

$r_2(\tau)$ : Autocorrelación de orden  $\tau$  de los cuadrados  $y_t^2$ .

$Q(20)$  y  $Q_2(20)$ : Estadístico Box Ljung para  $y_t$  e  $y_t^2$  respectivamente (31.4 es el valor crítico al 5%).

\* Significativamente distinto de cero al 5%.

Tabla 4.6: Modelos GARCH(1,1) gaussianos estimados para las series corregidas de atípicos siguiendo el procedimiento propuesto y diagnosis de los residuos

	US CAN	US ESP	US GER	US JAP	US SSF	US SWK	US UKP	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
$\hat{\omega}$	0.0009 (0.001)	0.0019 (0.001)	0.0029 (0.001)	0.0082 (0.002)	0.0061 (0.002)	0.0051 (0.002)	0.0047 (0.001)	0.0171 (0.006)	0.1031 (0.022)	0.0060 (0.002)	0.0320 (0.008)	0.0037 (0.001)
$\hat{\alpha}$	0.0461 (0.005)	0.0392 (0.006)	0.0382 (0.006)	0.0527 (0.006)	0.0314 (0.005)	0.0312 (0.006)	0.0374 (0.005)	0.0667 (0.009)	0.0990 (0.013)	0.0516 (0.006)	0.1073 (0.014)	0.0391 (0.004)
$\hat{\beta}$	0.9445 (0.006)	0.9569 (0.007)	0.9545 (0.008)	0.9333 (0.007)	0.9549 (0.009)	0.9545 (0.010)	0.9416 (0.008)	0.9242 (0.011)	0.8687 (0.016)	0.9416 (0.007)	0.8710 (0.018)	0.9553 (0.005)
$\hat{\alpha} + \hat{\beta}$	0.9906	0.9961	0.9927	0.9860	0.9863	0.9857	0.9790	0.9909	0.9677	0.9932	0.9783	0.9944
log L	342	-1732	1725	2086	1960	1787	1331	1984	2381	3337	2967	3364
$\hat{\varepsilon}_t$ $\frac{y_t}{\hat{\sigma}_t}$												
Media	0.0322	0.0433	0.0311	0.0067	0.0188	0.0274	0.0045	0.0134	0.0142	0.0056	0.0015	0.0077
D.T.	1.0013	1.0014	1.0012	1.0000	1.0012	1.0015	0.9986	1.0007	0.9999	1.0007	0.9999	1.0005
Simetría	0.0877	0.1593*	0.2409*	0.2250*	0.2603*	0.0624	0.0283	0.1128*	0.0308	0.1753*	0.1304*	0.0803
Curtosis	4.0586*	3.8501*	3.9639*	4.8491*	3.9185*	3.7281*	4.1632*	4.0356*	4.1797*	3.9632*	3.3970*	3.9077*
r(1)	0.0100	0.0360	0.0220	0.0060	0.0210	0.0050	0.0210	0.0090	0.0480	0.0320	0.0080	0.0050
Q(20)	24.85	18.85	18.95	26.95	19.81	15.88	27.66	23.37	30.90	27.54	20.99	26.23
$r_2(1)$	0.0360	0.0140	0.0070	0.0330	0.0260	0.0280	0.0320	0.0080	0.0270	0.0080	0.0060	0.0190
$r_2(2)$	0.0450	0.0170	0.0170	0.0030	0.0220	0.0250	0.0220	0.0450	0.0010	0.0110	0.0030	0.0010
$r_2(5)$	0.0150	0.0020	0.0030	0.0230	0.0090	0.0140	0.0140	0.0030	0.0030	0.0220	0.0090	0.0100
$r_2(10)$	0.0270	0.0170	0.0360	0.0310	0.0130	0.0080	0.0320	0.0430	0.0110	0.0040	0.0110	0.0050
Q <sub>2</sub> (20)	23.25	10.66	17.35	16.52	17.69	18.69	16.81	13.67	15.24	13.18	17.83	9.26





## Capítulo 5

# Detección de cambios de nivel

En este capítulo se propone un nuevo contraste para detectar cambios de nivel en series temporales. Se deriva la distribución asintótica del estadístico del contraste y, mediante experimentos de Monte Carlo se analizan sus propiedades en muestras finitas. El nuevo estadístico tiene la ventaja, frente a otros más tradicionales, de tener una distribución asintótica conocida que aproxima adecuadamente la distribución muestral. Además, el contraste es robusto a distribuciones no gaussianas y, en particular, a la presencia de heterocedasticidad condicional. En segundo lugar, se propone un procedimiento para contrastar múltiples cambios de nivel y, mediante un ejercicio de simulación, se muestran sus ventajas frente al procedimiento más utilizado en la actualidad. Finalmente, se aplica el contraste y el procedimiento propuestos a la serie de rendimientos diarios del tipo de cambio de la peseta frente al dólar estadounidense, en la que se detectan dos cambios de nivel.

## 5.1 Introducción

Cuando se analizan series temporales económicas con tamaños muestrales suficientemente grandes es común encontrar cambios de nivel, los cuales podrían estar causados por guerras, crisis financieras, huelgas, intervenciones políticas, etc. Este tipo de cambios produce sesgos importantes en la estimación de las funciones de autocorrelación y autocorrelación parcial de la serie y, como señalan Chen y Tiao (1990), hacen que los estimadores de mínimos cuadrados de los parámetros del modelo ARMA sean inconsistentes. Existen en la literatura varios estadísticos propuestos para identificar el punto de cambio en la media de un proceso estocástico. Por ejemplo, Hinkley (1971) propone estimar el punto de cambio en la media de una secuencia de variables normales usando un contraste tipo CUSUM. Worsley (1986) propone el uso de métodos de máxima verosimilitud para contrastar la presencia de un cambio en la media para una secuencia de variables independientes con distribución en la familia exponencial, estimar el punto de cambio y dar intervalos de confianza. Los métodos anteriores suponen independencia en el proceso, mientras que en un trabajo reciente, Kokoszka y Leipus (1998) prueban la consistencia de una familia de estimadores tipo CUSUM del punto de cambio en la media de una secuencia de observaciones dependientes y derivan las tasas de convergencia. No obstante, estos autores no abordan el problema de contrastar si existe un punto de cambio. Andrews (1993) considera diferentes estadísticos para contrastar la existencia de cambios estructurales en un instante desconocido. Los resultados se aplican a una amplia clase de modelos paramétricos que se pueden estimar por GMM (Método Generalizado de Momentos). Csorgo y Horvath (1997) es un libro de texto extenso en teoría asintótica para el análisis de puntos de cambio.

Tsay (1988) considera el problema de la detección de observaciones atípicas, cambios de nivel y cambios en la varianza en series temporales univariantes generadas por modelos ARMA con innovaciones gaussianas y su método es uno de los más utilizados

en la práctica. Sin embargo, Balke (1993) argumenta que dicho contraste puede no funcionar bien cuando la serie no es incorrelada porque, en presencia de cambios de nivel, las estimaciones del modelo ARMA serán sesgadas y, propone una modificación del estadístico de Tsay. Vogelsang (1998) aborda también el problema de la detección de cambios de nivel en series temporales, y propone un contraste que es robusto a la presencia de autocorrelación.

Dado que el tema central de esta Tesis, es la modelización de series financieras, que en la mayoría de los casos son series incorreladas con heterocedasticidad condicional, parece razonable estudiar el comportamiento del estadístico utilizado por Tsay cuando el supuesto de gaussianidad no es adecuado. En este capítulo, se analiza la distribución del estadístico usado por Tsay (1988), en tamaños de muestra finitos (su distribución asintótica no tiene una expresión analítica conocida) y se concluye que dicha distribución depende tanto de la distribución del proceso generador como del tamaño de la muestra. En consecuencia, la utilización empírica del estadístico de Tsay (1988) es limitada, dado que sería necesario generar valores críticos diferentes para distintos tamaños de muestra y/o distintas distribuciones.

Se propone un contraste alternativo basado en la idea de Inclan y Tiao (1994) de usar sumas acumuladas de cuadrados para la detección retrospectiva de cambios en la varianza de una secuencia de observaciones independientes. Como se verá a lo largo del capítulo, dicho contraste es robusto a la presencia de heterocedasticidad condicional y su distribución asintótica es una buena aproximación en muestras finitas. Se analiza también el problema de la detección de múltiples cambios de nivel en una serie. Se propone un procedimiento que consiste en detectar un cambio de nivel, dividir la serie en las dos subseries correspondientes y, buscar en ellas nuevos cambios. Se muestra que este procedimiento es mejor que el utilizado habitualmente, que consiste en detectar un cambio de nivel, corregir la serie por su magnitud y buscar otros cambios en la serie corregida. La ventaja fundamental del procedimiento propuesto es que su objetivo

es detectar todos los cambios existentes en la serie antes de estimar sus respectivos tamaños lo cual evita sesgos en la estimación y la generación de cambios de nivel espurios, que son los principales problemas del procedimiento más tradicional.

La organización de este capítulo es la siguiente. La sección 2 describe el contraste propuesto por Tsay (1988) e ilustra, mediante experimentos de Monte Carlo, sus problemas de tamaño y potencia cuando se implementa en muestras finitas. La sección 3 describe el nuevo estadístico propuesto y su relación con el estadístico de Tsay. Se deriva su distribución asintótica, se prueba su consistencia y se analizan sus propiedades en muestras finitas. La sección 4 aborda la detección de un cambio de nivel con autocorrelación en una serie y la sección 5 la detección de múltiples cambios de nivel. En la sección 6 se analiza la serie de rendimientos diarios del tipo de cambio peseta/dólar estadounidense, en la que se detectan dos cambios de nivel. Finalmente, las conclusiones se recogen en la sección 7.

## 5.2 Método de Tsay para detectar cambios de nivel

Considérese una serie temporal  $y_t$ ,  $t = 1, \dots, T$  y supóngase que la serie observada  $z_t$  viene dada por el modelo

$$z_t = y_t + wI(t \geq \tau) = \begin{cases} y_t + w & \text{si } t \geq \tau \\ y_t & \text{en otro caso} \end{cases} \quad (5.1)$$

donde  $w$  es el tamaño del cambio de nivel que tiene lugar en el instante  $\tau$ . Cuando el proceso generador de los datos  $y_t$ , es un ruido blanco<sup>1</sup>, el estadístico propuesto por

---

<sup>1</sup>Dado que la inmensa mayoría de las series reales de interés se caracterizan por tener dependencia temporal, Tsay (1988) propone modelizar la media condicional mediante un modelo ARIMA y aplicar el contraste a los residuos del modelo, que ya son ruido blanco. No obstante, si hay un cambio de nivel en la serie observada, los parámetros del modelo se estimarán con sesgos y, por tanto, este procedimiento puede ser inadecuado. En la sección 4 se tratará este problema más detalladamente.

Chen y Tiao (1986) y generalizado por Tsay (1988) para contrastar la hipótesis nula  $H_0 : w = 0$ , viene dado por  $\lambda = \max_{m=1, \dots, T} |\lambda_m|$ , donde

$$\lambda_m = \frac{\sum_{i=m+1}^T (z_i - \bar{z})}{\hat{\sigma} \sqrt{\frac{m(T-m)}{T}}} \quad m = 1, 2, \dots, T-1 \quad (5.2)$$

con  $\bar{z} = \sum_{i=1}^T z_i / T$ ,

$$\hat{\sigma}^2 = \hat{\sigma}_z^2 - \frac{1}{m(T-m)} \left[ \sum_{i=m+1}^T (z_i - \bar{z}) \right]^2 \quad (5.3)$$

y  $\hat{\sigma}_z = \sqrt{\frac{\sum_{i=1}^T (z_i - \bar{z})^2}{T}}$  (véase el Apéndice A para más detalles). Alternativamente,  $\lambda_m$  se puede reescribir como

$$\lambda_m = \frac{\bar{z}_2 - \bar{z}_1}{\frac{\hat{\sigma}}{\sqrt{\frac{m(T-m)}{T}}}} \quad m = 1, 2, \dots, T-1 \quad (5.4)$$

donde  $\bar{z}_1 = \sum_{i=1}^m z_i / m$  y  $\bar{z}_2 = \sum_{i=m+1}^T z_i / (T-m)$  son las medias muestrales antes y después del instante  $m$  respectivamente. Es inmediato ver que si  $y_t$  es un proceso gaussiano, bajo la hipótesis nula  $H_0 : w = 0$ ,  $\lambda_m$  se distribuye según una  $N(0, 1)$ , para todo  $m$ . Nótese que, a pesar de que la distribución asintótica de  $\lambda_m$  es conocida, para todo  $m$ , la distribución asintótica del estadístico  $\lambda$ , si existe, es desconocida. El estadístico  $\lambda_m$  puede interpretarse como una diferencia ponderada entre la media de las  $m$  primeras observaciones y la media de las  $T-m$  últimas observaciones de la serie. Cuando dicha diferencia supere un determinado valor crítico, la hipótesis nula de que la media es la misma, o equivalentemente, que  $w = 0$  se rechazará. Tsay (1988) propone comparar el estadístico del contraste,  $\lambda$ , con un valor crítico: 3.5, 3 o 2.5, para lo que se basa en un ejercicio de simulación llevado a cabo para 5000 réplicas de procesos ruido blanco gaussianos y modelos AR(1). Si  $\lambda$  es mayor que el valor crítico elegido, el momento  $t^*$ , tal que  $\lambda = |\lambda_{t^*}| = \max_{m=1, \dots, T} |\lambda_m|$ , es identificado como el instante en el que tiene lugar un cambio de nivel.

En la tabla 5.1 aparecen los percentiles empíricos del estadístico  $\lambda$  basados en 10000 réplicas de observaciones independientes generadas por una distribución  $N(0, 1)$



con distintos tamaños de muestra,  $T = 25, 200, 500, 1000, 5000$  y  $15000$ . La figura 5.1 representa estimaciones kernel de la densidad de dicho estadístico para tres tamaños muestrales,  $T = 25, 200$  y  $15000$ . Obsérvese que la distribución del estadístico  $\lambda$  depende claramente del tamaño muestral, y por ello no es posible elegir un valor crítico adecuado para cualquier tamaño muestral. Por ejemplo, para un nivel de significación de 10%, el valor crítico sería 2.98 cuando  $T = 25$  y 3.18 cuando  $T = 15000$ . Por otra parte, el mismo valor crítico, por ejemplo, 3.01, corresponde a un nivel de significación de 15% para  $T = 5000$  y a un nivel de 10% para  $T = 500$ . Los valores críticos sugeridos por Tsay (1988), 3.5, 3 o 2.5 podrían, por tanto, corresponder a niveles de significación de 5%, 15% y 25% respectivamente para tamaños de muestra grandes. Los valores críticos que se usarán en lo que sigue son los percentiles 90, 95 y 99 de la distribución empírica cuando el tamaño muestral es  $T = 15000$ , que aproximarían los percentiles de la distribución asintótica si ésta existe, y que son 3.18, 3.43 y 3.90 como se puede observar en la tabla 5.1.

Tsay (1988) propone el estadístico  $\lambda$  para series gaussianas. Como, en general, las series financieras no son gaussianas, a continuación se va a analizar si este contraste es robusto a la presencia de heterocedasticidad condicional. Con este objetivo se ha calculado el tamaño del contraste cuando se aplica a series generadas por modelos de heterocedasticidad condicional. En particular, se han generado 10000 réplicas de tamaños  $T = 500, 1000$  y  $5000$ , mediante cuatro modelos diferentes: (i) GARCH(1,1) gaussiano, (ii) GARCH(1,1)- $t_7$ , (iii) EGARCH(1,1) gaussiano, y (iv) ARSV(1) gaussiano. La tabla 5.2 muestra el tamaño cuando se consideran los valores críticos  $C_T = 3.18, 3.43$  y  $3.90$  correspondientes a niveles de significación de 10%, 5% y 1% respectivamente. A pesar de que se esperaría que el tamaño empírico del contraste fuese menor que el tamaño nominal, ya que se están considerando valores críticos mayores que los que corresponden al tamaño muestral, se puede observar que el tamaño empírico es, en la mayoría de los casos, superior al nominal. Esto es exagerado en algunos casos como, por

ejemplo, para los cuatro modelos ARSV(1), donde algunos tamaños empíricos doblan los nominales. Además, es muy importante señalar que, en general, a medida que el tamaño muestral aumenta, los tamaños empíricos se hacen mayores alejándose en muchos casos de los nominales. Por lo tanto, se deduce que las propiedades de este contraste en lo que a tamaño se refiere no son adecuadas. El contraste no es robusto ante ciertas formas de heterocedasticidad condicional.

Las tablas 5.3 y 5.4 contienen la potencia del contraste. Para ello, se han generado 5000 réplicas de tamaños muestrales  $T = 500, 1000$  y  $5000$  bajo la hipótesis alternativa, esto es, siguiendo el modelo (5.1) para distintos valores de  $w$ . En la tabla 5.3, la serie,  $y_t$ , sigue una distribución  $N(0,1)$ , y en la tabla 5.4,  $y_t$  se ha generado a partir del modelo GARCH(1,1) gaussiano con parámetros  $(\omega, \alpha, \beta) = (0.07, 0.05, 0.88)$ . Se han considerado además tres instantes distintos para el cambio de nivel:  $\tau = T/10, T/2$  y  $9T/10$ . Para que las potencias sean comparables se han elegido los valores críticos para que en cada tamaño muestral, el tamaño del contraste sea 5%. Como se puede observar en ambas tablas, el contraste tiene buenas propiedades en lo que a potencia se refiere. Además, como es de esperar, la potencia del contraste aumenta con el tamaño del cambio y con el tamaño muestral. Nótese también que la potencia es mayor cuando el cambio ocurre en el centro de la serie ( $\tau = T/2$ ). Por ejemplo, para una muestra de tamaño  $T = 5000$ , un cambio de nivel de tamaño  $w = 0.2$  desviaciones típicas, se detecta con probabilidad 1 cuando tiene lugar en el centro de la serie, y con probabilidad 0.9, aproximadamente, cuando ocurre en los extremos. Los resultados no parecen depender en gran medida del modelo con el que se han simulado los datos, aunque la potencia del contraste es ligeramente menor cuando la serie se ha generado con el modelo GARCH.

### 5.3 Un contraste alternativo. Distribución asintótica y propiedades en muestras finitas

En esta sección se propone un contraste nuevo para detectar cambios de nivel y se deriva la distribución asintótica del estadístico propuesto. Considérese de nuevo la serie temporal  $y_t$ ,  $t = 1, \dots, T$  y supóngase que la serie observada  $z_t$  viene dada por (5.1). El estadístico que se propone en este capítulo viene dado por  $e = \max_{m=1, \dots, T} |e_m|$  donde

$$e_m = \frac{\sum_{i=1}^m z_i - \frac{m}{T} \sum_{i=1}^T z_i}{\sqrt{\sum_{i=1}^T (z_i - \bar{z})^2}}. \quad (5.5)$$

Bajo la hipótesis nula  $H_0 : w = 0$ , la distribución asintótica del estadístico  $e$  se deriva en el siguiente teorema.

**Teorema 1** *Sea  $\{y_1, y_2, \dots, y_T\}$  un proceso estacionario y ergódico para el cual*

$$E(y_t | y_1, \dots, y_{t-1}) = 0$$

*con probabilidad 1 y  $E(y_t^2) = \sigma_y^2$  es positiva y finita. Sea  $z_t$  el proceso dado por (5.1) y sea  $e_m$  dado por (5.5). Entonces, bajo la hipótesis nula  $H_0 : w = 0$ , la distribución asintótica del estadístico  $e = \max_{m=1, \dots, T} |e_m|$  viene dada por la distribución del supremo de un puente Browniano, es decir,*

$$P(e \leq x) \longrightarrow G(x)$$

donde

$$G(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - 2 \sum_{k=1}^{\infty} (-1)^{k+1} e^{-2k^2 x^2} & \text{si } x \geq 0 \end{cases}$$

**Demostración:** Sea  $X_T(j) = \frac{1}{\sigma_z \sqrt{T}} \sum_{i=1}^{[Tj]} z_i$ , donde  $[Tj]$  es el mayor entero menor o igual que  $Tj$  y  $j \in [0, 1]$ . Aplicando el Teorema 23.1 en Billingsley (1968) se tiene que

$$X_T \longrightarrow B \quad \text{débilmente} \quad (5.6)$$



donde  $B$  es un movimiento Browniano. Por el teorema de la aplicación continua se obtiene que

$$X_T(j) - jX_T(1) \longrightarrow B^0 \quad \text{débilmente} \quad (5.7)$$

donde  $B^0$  es un puente Browniano. Sea  $m = Tj$  para  $m = 1, 2, \dots$ . Entonces,  $X_T(j) - jX_T(1) = \frac{1}{\sigma_z \sqrt{T}} \sum_{i=1}^m z_i - \frac{m}{T} \left( \frac{1}{\sigma_z \sqrt{T}} \sum_{i=1}^T z_i \right) = \frac{1}{\sigma_z \sqrt{T}} \left\{ \sum_{i=1}^m z_i - \frac{m}{T} \sum_{i=1}^T z_i \right\}$ . Puesto que  $\frac{1}{T} \sum_{i=1}^T (z_i - \bar{z})^2 \longrightarrow \sigma_z^2$  se tiene  $e_m = \frac{\sum_{i=1}^m z_i - \frac{m}{T} \sum_{i=1}^T z_i}{\sqrt{\sum_{i=1}^T (z_i - \bar{z})^2}} \longrightarrow B^0$ . Sabiendo que  $e = \max_{m=1, \dots, T} |e_m|$  y que  $e_m \longrightarrow B^0$  débilmente, a partir de Kennedy (1976) se obtiene que

$$P(e \leq x) \longrightarrow G(x) \blacksquare$$

La tabla 5.5 contiene los percentiles de la distribución asintótica y de la distribución empírica del estadístico  $e$ , calculados a partir de 10000 réplicas de series independientes  $N(0,1)$  para los mismos tamaños muestrales considerados en la sección anterior. La figura 5.2 muestra estimadores kernel de la densidad del estadístico para  $T = 25, 200$  y 15000. Si se compara este gráfico con la figura 5.1, puede observarse que la función de densidad del estadístico  $e$  es más estable que la del estadístico  $\lambda$ , en el sentido de que la distribución asintótica se alcanza relativamente rápido. Por ejemplo, para un tamaño de muestra  $T = 200$ , la aproximación asintótica es ya buena. Por lo tanto, el estadístico  $e$ , depende en menor medida del tamaño muestral que el estadístico de Tsay,  $\lambda$ .

No obstante, la mayor ventaja de este procedimiento es que se conoce la distribución asintótica del proceso, mientras que en el contraste de Tsay se conoce solamente la distribución de  $\lambda_m$  para cada  $m$ . Además, para derivar la distribución de  $\lambda_m$  se necesita el supuesto de normalidad, mientras que la distribución de  $e$  se obtiene sin hacer supuestos distribucionales para la serie observada  $z_t$ .

Es posible establecer una relación entre los estadísticos  $\lambda$  y  $e$  que puede ayudar a

clarificar su comportamiento en la práctica. Dado que  $e_m$  se puede reescribir como

$$e_m = \frac{\bar{z}_1 - \bar{z}_2}{\frac{\hat{\sigma}_z}{\sqrt{\frac{m(T-m)}{T}}}} \frac{1}{\sqrt{\frac{m(T-m)}{T}}} \quad (5.8)$$

es fácil ver que el proceso  $e_m$  se puede escribir en función de  $\lambda_m$  del siguiente modo:

$$e_m = -\frac{\sqrt{m(T-m)}}{T} \frac{\hat{\sigma}}{\hat{\sigma}_z} \lambda_m \quad m = 1, 2, \dots, T-1. \quad (5.9)$$

con  $\hat{\sigma}$  dada por (5.3). Por lo tanto, el estadístico propuesto en este capítulo puede interpretarse como el estadístico de Tsay corregido de modo tal que tiene una distribución asintótica conocida. Como  $\hat{\sigma} < \hat{\sigma}_z$ , en (5.9) se tiene que  $\frac{\sqrt{m(T-m)}}{T} \frac{\hat{\sigma}}{\hat{\sigma}_z} < 1$ , entonces  $\lambda_m$  será siempre mayor, en valor absoluto, que  $e_m$  y consecuentemente, el estadístico  $\lambda$  siempre será mayor que  $e$ .

En resumen, el procedimiento que se propone seguir para contrastar la hipótesis nula  $H_0 : w = 0$  en la serie temporal observada,  $z_1, z_2, \dots, z_T$  consiste en: (i) calcular  $e = \max_{m=1, \dots, T} |e_m|$  y (ii) una vez fijado el valor crítico  $C_N$  de la tabla 5.5 al nivel de confianza elegido, si  $e = |e_{t^*}| > C_N$ , hay un cambio de nivel en el instante  $t^*$ .

En el siguiente teorema se prueba la consistencia del contraste propuesto.

**Teorema 2** Sea  $\{y_1, y_2, \dots, y_T\}$  bajo los mismos supuestos del Teorema 1. Bajo la hipótesis alternativa  $H_A : w \neq 0$ ,

$$\lim_{T \rightarrow \infty} P(e > C_N) = 1.$$

donde  $C_N$  es el valor crítico.

**Demostración:** Dado que  $z_t = y_t + wI(t > \tau)$  y  $w \neq 0$ , se tiene

$$e_m^{H_A} = \frac{\sum_{i=1}^m z_i - \frac{m}{T} \sum_{i=1}^T z_i}{\sqrt{\sum_{i=1}^T (z_i - \bar{z})^2}}.$$

Como  $\sum_{i=1}^T z_i = \sum_{i=1}^T y_i + (T - \tau)w$ ,  $\bar{z} = \bar{y} + \frac{T-\tau}{T}w$  y

$$\sum_{i=1}^T (z_i - \bar{z})^2 = \sum_{i=1}^T (y_i - \bar{y})^2 + \frac{w^2 \tau (T - \tau)}{T} - 2w \sum_{i=1}^{\tau-1} y_i + \frac{2w\tau}{T} \sum_{i=1}^T y_i$$

Entonces, denotando por  $e_m^{H_A}$  al valor del estadístico del contraste bajo la hipótesis alternativa, se tiene

$$e_m^{H_A} = \begin{cases} \frac{\sum_{i=1}^m y_i - \frac{m}{T} \sum_{i=1}^T y_i - \frac{m(T-\tau)w}{T}}{\sum_{i=1}^T (y_i - \bar{y})^2 + \frac{w^2 \tau (T-\tau)}{T} - 2w \sum_{i=1}^{\tau-1} y_i + \frac{2w\tau}{T} \sum_{i=1}^T y_i} & \text{si } m < \tau \\ \frac{\sum_{i=1}^m y_i - \frac{m}{T} \sum_{i=1}^T y_i + (m-\tau)w - \frac{m(T-\tau)w}{T}}{\sum_{i=1}^T (y_i - \bar{y})^2 + \frac{w^2 \tau (T-\tau)}{T} - 2w \sum_{i=1}^{\tau-1} y_i + \frac{2w\tau}{T} \sum_{i=1}^T y_i} & \text{si } m \geq \tau \end{cases} \quad (5.10)$$

Utilizando un supuesto estándar; véase, por ejemplo, Vogelsang (1998), y es suponer que  $\tau = kT$ , donde  $0 < k < 1$  es una constante, es decir,  $\frac{\tau}{T}$  permanece constante a medida que el tamaño muestral aumenta, es fácil ver que

$$\lim_{T \rightarrow \infty} e_m^{H_A} = \pm \infty \quad (5.11)$$

dependiendo del signo de  $w$ . Si  $w > 0$ , el límite en (5.11) será  $-\infty$ , mientras que si  $w < 0$ , dicho límite será  $+\infty$ . Por lo tanto, la probabilidad de rechazar la hipótesis nula cuando se está bajo la alternativa, es decir,  $P(e > C_N)$ , tiende a 1 a medida que el tamaño muestral,  $T$ , tiende a  $\infty$  ■

Nótese que la derivación de la distribución asintótica del estadístico  $e$  sólo requiere que la serie  $y_t$  sea incorrelada, estacionaria y ergódica con varianza finita. Por lo tanto, el contraste puede utilizarse en series no gaussianas y, en particular, en series condicionalmente heterocedásticas. La tabla 5.6 contiene el tamaño del contraste propuesto, para las mismas series condicionalmente heterocedásticas de la tabla 5.2 y considerando tres valores críticos diferentes:  $C_N = 1.22, 1.36$  y  $1.63$ , correspondientes a niveles de significación de 10%, 5% y 1% respectivamente. Se puede observar que el tamaño del nuevo contraste es, en general, muy cercano al nominal, excepto en el caso del modelo GARCH- $t_7$  con parámetros  $(\omega, \alpha, \beta) = (0.02, 0.10, 0.88)$ , para el cual el tamaño

empírico es muy superior al nominal, al igual que lo que sucedía con el contraste Tsay. En este caso los parámetros del modelo no satisfacen la condición de existencia de la curtosis y por tanto la estimación de la varianza puede no ser precisa lo que podría llevar a rechazar la hipótesis nula más veces de las debidas. No obstante, las propiedades del nuevo contraste en términos de tamaño mejoran considerablemente las de Tsay, especialmente en los modelos ARSV. El nuevo contraste es robusto a distribuciones no gaussianas y en particular a la presencia de heterocedasticidad condicional.

La potencia del contraste se muestra en las tablas 5.7 y 5.8. Al igual que en las tablas 5.3 y 5.4, se han generado 5000 réplicas de tamaños muestrales  $T = 500, 1000$  y 5000 bajo la hipótesis alternativa para distintos valores de  $w$  y para distintos instantes en los que ocurre el cambio de nivel. El tamaño del contraste se ha fijado al 5% eligiendo en cada caso el valor crítico adecuado. Los resultados que se obtienen son similares a los mostrados para el contraste de Tsay. Así, por ejemplo, nótese que la potencia del contraste aumenta con el tamaño muestral y con la magnitud del cambio,  $w$ . Dicha potencia también es mayor cuando el cambio ocurre en el centro de la serie que cuando ocurre en los extremos. Si se compara la potencia del nuevo contraste con la del contraste de Tsay se observa que éste último tiene más potencia cuando el cambio tiene lugar en los extremos de la serie. Sin embargo, cuando el cambio ocurre en el centro de la serie, el nuevo contraste tiene mejores propiedades. De nuevo, los resultados parecen ser independientes del modelo con el que se han generado los datos. Cuando el proceso generador no es un proceso gaussiano sino GARCH, no parece haber pérdida de potencia, mientras que con el contraste de Tsay había una ligera pérdida.

## 5.4 Detección de un cambio de nivel en modelos ARIMA

Las series de rendimientos financieros son, en la mayoría de los casos, incorreladas y consecuentemente, el contraste propuesto en este capítulo podría aplicarse a este tipo de series tal y como se ha descrito en la sección anterior. No obstante, hay otras muchas series económicas y financieras de interés que están correladas, y por tanto ya no satisfacen los supuestos del teorema 1. Cuando hay correlación entre las observaciones, Tsay propone ajustar, en primer lugar, un modelo ARIMA a la serie observada y después aplicar el contraste a los residuos del modelo estimado. No obstante, como nota Balke (1993), la presencia de cambios de nivel en la serie puede llevar a estimar con sesgos los parámetros del modelo. Balke (1993) propone una modificación del contraste de Tsay para solucionar este problema. Vaage (2000) estudia las propiedades del contraste propuesto por Balke (1993) y concluye que es muy sensible a la presencia de cambios de nivel transitorios, que tiende a identificar como permanentes, pero que, en general, el procedimiento funciona mejor que el propuesto por Tsay. Alternativamente, Vogelsang (1998) propone un contraste para detectar un cambio en la media de series temporales univariantes, que no requiere la estimación de parámetros. Vogelsang (1999) estudia la potencia de varios contrastes para detectar cambios de nivel en series temporales, encontrando distintas razones que hacen que muchas de las curvas de potencia sean no monótonas.

Para analizar el comportamiento del estadístico e cuando se aplica a los residuos de un modelo estimado se han realizado varios experimentos de Monte Carlo. Se han simulado 10000 réplicas de tres modelos autorregresivos estacionarios con innovaciones gaussianas: (i) AR(1) con parámetro  $\phi = 0.5$ , (ii) AR(1) con parámetro  $\phi = 0.9$  y finalmente (iii) AR(3) con parámetros  $\phi_1 = 0.2$ ,  $\phi_2 = 0.5$  y  $\phi_3 = -0.1$ . La tabla 5.9 contiene el ratio, en porcentaje, entre los percentiles de la distribución empírica de e

cuando se aplican a los residuos de los tres modelos estimados y los percentiles de la tabla 5.5. Como se puede observar, en general, los resultados son muy próximos a 100 sugiriendo por tanto, que ambas distribuciones empíricas son muy similares y que la teoría asintótica podría ser la misma en el caso de utilizar residuos. No obstante, este resultado se basa en simulaciones por lo que no se puede asegurar que sea cierto, pero será objeto de investigación futura. Otro problema a solventar sería el mencionado por Balke (1993), es decir, como tratar los sesgos en las estimaciones de los parámetros que pueden venir causados por la existencia de un cambio de nivel.

## 5.5 Detección de múltiples cambios de nivel

Tanto el contraste de Tsay como el que se propone en este capítulo están diseñados para detectar un único cambio de nivel. Sin embargo, pueden existir varios cambios en una misma serie. En ese caso, Tsay (1988) propone detectar el instante en que ocurre el mayor cambio, estimar su tamaño,  $w$ , corregir la serie y volver a aplicar el procedimiento a la serie corregida, repitiendo el procedimiento hasta que el contraste no detecte ningún cambio. Es inmediato pensar que, la estimación del tamaño del primer cambio de nivel detectado va a estar sesgada por la presencia de los demás. Una alternativa a este procedimiento, que evitaría estimar en cada etapa el tamaño del cambio, es dividir la serie en dos subseries: una hasta el instante en que se detecta el primer cambio de nivel y la otra, a partir del mismo, y aplicar el contraste a las subseries suponiendo que son series diferentes. En concreto, esta alternativa es:

**Paso 0** Estimar, si hay estructura en la media de la serie, el proceso ARIMA correspondiente como en Tsay (1988), y quedarse con los residuos que ya son incorrelados.

**Paso 1** Calcular  $e$ . Fijando el valor crítico  $C_N$  de la tabla 5.5 al nivel de confianza deseado:

- Si  $e < C_N$ : No hay cambios de nivel en la serie observada.
- Si  $e > C_N$ : Hay un cambio de nivel en el instante correspondiente. Ir al paso siguiente.

**Paso 2** Denótese por  $t^*$  el instante en que se alcanza el máximo, es decir,  $e = |e_{t^*}| = \max_{m=1, \dots, T} |e_m|$ . Considerar las subseries  $y_1, y_2, \dots, y_{t^*}$  e  $y_{t^*+1}, \dots, y_T$  y repetir los Pasos 0 y 1 para cada una de las subseries. Continuar hasta que no se detecte ningún cambio.

Una posible debilidad de este procedimiento es que, cuando la serie se divide, los tamaños muestrales de las subseries disminuyen lo que hace que la potencia del contraste también disminuya. Además, en ambas subseries se usa el mismo valor crítico a pesar de que éstos se basan en resultados asintóticos. No obstante, cuando se trabaja con series financieras, los tamaños muestrales son tan grandes que se esperaría que el método funcionase bien. Bai (1997) propone, usando un estadístico diferente, utilizar este procedimiento para detectar más de un cambio de nivel en un proceso lineal.

A continuación, se comparan ambos procedimientos que se denotarán por, **(D)**: detectar un cambio de nivel, dividir la serie en las dos subseries correspondientes y buscar nuevos cambios en las subseries, y **(C)**: detectar un cambio de nivel, corregir la serie por su magnitud y repetir el procedimiento. Se han simulado 10000 réplicas de series independientes  $N(0, 1)$ , con tamaños muestrales  $T = 500$  y  $1000$ , contaminadas con dos cambios de nivel del mismo tamaño  $w = 1$ , variando la posición en la que ocurren en la serie: al principio ( $\tau = T/10$  o  $\tau = T/4$ ), en el centro ( $\tau = T/2$ ) o al final ( $\tau = 3T/4$  o  $\tau = 8T/10$ ) y variando la separación entre los cambios: que ambos estén relativamente juntos (la separación es  $T/10$ ) o separados (la separación es  $T/4$ ).

La serie contaminada viene dada por

$$z_t = y_t + wI(t \geq \tau_1) + wI(t \geq \tau_2) = \begin{cases} y_t & \text{si } t < \tau_1 \\ y_t + w & \text{si } \tau_1 \leq t < \tau_2 \\ y_t + 2w & \text{si } t \geq \tau_2 \end{cases} \quad (5.12)$$

Las tablas 5.10 y 5.11 contienen los resultados del experimento, para  $T = 500$  y  $T = 1000$  respectivamente, cuando el contraste utilizado es el propuesto en este capítulo, que usa el estadístico  $e$ . La primera columna contiene el porcentaje de veces que el contraste rechaza la hipótesis nula en la serie observada, las dos columnas siguientes contienen, en el procedimiento **(D)**, el porcentaje de veces que se rechaza la hipótesis nula cuando el contraste se aplica a las subseries obtenidas diviendo la serie original en dos subseries. En el procedimiento **(C)**, las columnas 2 y 3 contienen el porcentaje de rechazos de  $H_0$  cuando el contraste se aplica a las series corregidas, después de estimar el tamaño del cambio. En cuanto a las tres últimas columnas, éstas muestran la observación mediana detectada como instante en que ocurre el cambio. Las conclusiones más importantes que se extraen de ambas tablas son:

- El procedimiento **(D)** parece funcionar mejor que el **(C)**, pues éste último implica estimar el tamaño del cambio de nivel en cada etapa, introduciendo con ello un sesgo en cada estimación. Las tablas 5.10 y 5.11 ponen de manifiesto el problema. Observando, por ejemplo, la segunda fila, es claro que el procedimiento **(D)** identifica dos cambios de nivel en las observaciones, prácticamente, exactas. Sin embargo, el procedimiento **(C)** detecta un tercer cambio espurio, que sería consecuencia de haber corregido la serie, erróneamente. En las filas 1, 6 y 7 se observa el mismo fenómeno. La figura 5.3 ilustra este problema. Como se puede observar en el gráfico, se tiene una serie de tamaño muestral 500,  $y_t \geq 0, \forall t$  que ha sido contaminada con dos cambios de tamaño  $w = 1$ , en las observaciones 50 y 100<sup>2</sup>. Aplicando el procedimiento **(C)**, no se llega a recuperar la serie  $y_t$ ,

---

<sup>2</sup>Esta serie puede interpretarse como la media de la serie  $z_t$  dada por (5.12) con  $y_t$  una serie de



pues en cada etapa se estima, erróneamente, el tamaño del cambio y, por tanto, al corregir la serie, se introduce un nuevo cambio. Con el procedimiento (D), el tamaño de los cambios se estima una sola vez.

- El primer cambio detectado parece ser el que tiene lugar en  $\tau_2$ , cuando los cambios ocurren al principio de la serie y el que ocurre en  $\tau_1$  cuando los cambios ocurren al final de la serie. Esto es lo esperable una vez que se analiza el estadístico  $e$ . Considérese la serie del ejemplo anterior en la que  $\tau_1 = T/10 = 50$  y  $\tau_2 = T/10 + T/10 = 100$ . La figura 5.4 contiene los valores de  $e_m$  para la serie  $z_t$  correspondiente. Denotando por  $s = \sqrt{\sum_{i=1}^T (z_i - \bar{z})^2}$ , es fácil probar que

$$e_{T/10} = -\frac{17T}{100s}, \quad e_{2T/10} = \frac{24T}{100s}$$

y por tanto, el máximo se alcanza en  $\tau_2$ . Cuando los cambios tienen lugar en el centro de la serie, es decir,  $\tau_1 = T/2 = 250$  y  $\tau_2 = T/2 + T/10 = 300$ , se puede probar igualmente que

$$e_{T/2} = \frac{9T}{20s}, \quad e_{T/2+T/10} = -\frac{8.8T}{20s}$$

La figura 5.4 contiene los valores de  $e_m$ . La diferencia entre  $e_{T/2}$  y  $e_{T/2+T/10}$  es muy pequeña pero se tiene que el máximo se alcanza en  $\tau_1$ . Esto implica que la observación que debería detectarse como primer cambio es  $t^* = T/2$ . Sin embargo, en las tablas 5.10 y 5.11 puede observarse que la observación mediana que se detecta como primer cambio es la observación 259, cuando el tamaño muestral es 500 y 513 cuando el tamaño muestral es 1000. Esto se debe a la variabilidad en la muestra. Véase la figura 5.5. No obstante, el primer cambio se detecta más próximo a  $\tau_1$  que a  $\tau_2$ . Cuando los cambios ocurren al final de la muestra, el primer cambio se detecta en  $\tau_1$ , por la misma razón por la que se detecta en  $\tau_2$  cuando los cambios ocurren al principio.

- El procedimiento **(D)** parece funcionar mejor cuando los cambios de nivel están separados que cuando están juntos y cuando ocurren en los extremos de la serie que cuando tienen lugar en el centro de la serie. Como se ha dicho anteriormente, en este último caso, el instante en que se detecta el cambio suele ser incorrecto debido a que los valores de  $e_m$  en los puntos intermedios, es decir,  $\tau_1 \leq m \leq \tau_2$ , son muy próximos.

Estos resultados con respecto a los procedimientos **(C)** y **(D)** parecen ser independientes del contraste utilizado. Como ilustración se ha repetido el experimento anterior con el estadístico usado por Tsay,  $\lambda$ . Los resultados se muestran en las tablas 5.12 y 5.13. Las conclusiones son las mismas que las obtenidas cuando el contraste se realiza con el estadístico  $\mathbf{e}$ , es decir, el procedimiento **(C)** funciona peor porque genera cambios espurios. De ambas tablas se puede concluir además que el contraste propuesto en este capítulo funciona mejor que el usado por Tsay, a la hora de identificar el punto de cambio. Obsérvese que la mediana de las observaciones que se detectan, con el estadístico  $\mathbf{e}$ , como el instante en que tiene lugar el cambio de nivel, están más cerca de las verdaderas que las detectadas con el estadístico  $\lambda$ .

## 5.6 Aplicación empírica

Este apartado ilustra el procedimiento propuesto analizando la serie de los rendimientos del tipo de cambio entre la peseta frente al dólar estadounidense<sup>3</sup>, observados diariamente, desde el 2 de Enero de 1980 al 18 de Abril de 2001. La serie de interés será la serie de rendimientos definida como  $y_t = 100(\log(p_t) - \log(p_{t-1}))$  donde  $p_t$  es el número de pesetas por dólar en el instante  $t$ , y el tamaño muestral es  $T = 5371$ . La figura 5.6 contiene la serie original  $p_t$  y la serie de los rendimientos  $y_t$ , junto con estimadores

---

<sup>3</sup>Al igual que los tipos de cambio analizados en los capítulos anteriores, los datos han sido obtenidos de la página web <http://pacific.commerce.ubc.ca/xr/>.

kernel de la densidad y el correlograma de los rendimientos al cuadrado. La tabla 5.14 muestra estadísticos descriptivos de la serie  $y_t$ . Como se ha visto a lo largo de la Tesis con otras series financieras, se observa un elevado coeficiente de curtosis y autocorrelaciones de las observaciones al cuadrado significativas.

Dado que la serie de rendimientos es incorrelada, no es necesario ajustar la media, de modo que se aplica el procedimiento propuesto a la serie  $y_t$ . La figura 5.7 representa los valores de  $e_m$  cuyo máximo,  $e$ , corresponde a la observación  $t^* = 1289$  y es 1.6020. Puesto que  $e > C_N = 1.36$ , el contraste rechaza, al 5%, la hipótesis nula, indicando, en el instante  $t = 1289$  que corresponde al 26 de febrero de 1985, la presencia de un cambio de nivel en la serie de rendimientos, o lo que es lo mismo, un cambio en la pendiente de la serie original de precios,  $p_t$ . A continuación se divide la serie original en dos subseries  $y_{1t}$  e  $y_{2t}$  y se aplica el procedimiento de nuevo, a cada una de ellas. Como ilustra la figura 5.7, el contraste detecta un nuevo cambio de nivel en el instante  $t = 3189$  que corresponde al día 2 de septiembre de 1992. Si se aplica de nuevo el procedimiento a las tres nuevas subseries: la primera desde  $t = 1$  hasta  $t = 1289$ , la segunda desde  $t = 1290$  hasta  $t = 3189$  y la última desde  $t = 3190$  hasta  $t = 5371$ , ya no se detectan más cambios de nivel. Como se puede observar en la tabla 5.14, estas tres subseries tienen medias diferentes.

Aplicando el contraste de Tsay a la serie  $y_t$ , se detecta un cambio de nivel en el instante  $t = 1289$ , al igual que con el contraste propuesto. Una vez detectado el cambio, se estima el tamaño del cambio, se corrige la serie por el mismo y se vuelve a contrastar la presencia de otro cambio. Cuando se aplica el contraste a la serie corregida, ya no aparecen más cambios de nivel, como se observa en la figura 5.8. Por lo tanto, el contraste de Tsay detecta solamente un cambio de nivel mientras que el procedimiento propuesto en este capítulo detecta dos cambios de nivel.

Se ha aplicado también el procedimiento (D) con el estadístico  $\lambda$  y el procedimiento (C) con el estadístico  $e$ . En el primer caso, el estadístico  $\lambda$  detecta un nuevo cambio

en la observación  $t = 1290$ , esto es, en la primera observación de la segunda subserie. Si se aplica el contraste de nuevo a la subserie  $y_{2t}$  para  $t = 1291, \dots, 5371$  ya no se detectan más cambios. En el segundo caso, el estadístico  $e$  detecta un cambio en la observación  $t = 3189$ . Por tanto, para este ejemplo concreto, independientemente del procedimiento usado, el estadístico  $e$  detecta dos cambios de nivel en la serie.

Es importante señalar que, aunque el estadístico  $\lambda$  sólo detecta un cambio, en realidad ocurrieron los dos cambios de nivel que el estadístico  $e$  ha detectado. En 1985 el Grupo de los 5 países más industrializados del mundo decidieron, en Washington D.C., debilitar el dólar frente a otras monedas del mundo con el objetivo de fomentar las exportaciones. Y en Septiembre de 1992 tuvo lugar una devaluación de la peseta.

## 5.7 Conclusiones

En este capítulo se han estudiado las propiedades del estadístico propuesto por Chen y Tiao (1986) y generalizado por Tsay (1988), comúnmente usado para detectar cambios de nivel en series temporales. La distribución asintótica de dicho estadístico no es conocida y se ha visto que su distribución en muestras finitas depende crucialmente del tamaño muestral y de la distribución del proceso generador de los datos. Se ha propuesto un nuevo contraste, para el que se ha encontrado su distribución asintótica. Mediante experimentos de Monte Carlo se han ilustrado sus propiedades en muestras finitas y mostrado que la distribución asintótica proporciona una buena aproximación de dichas propiedades. Se han comparado también dos procedimientos para detectar múltiples cambios de nivel en una misma serie y se ha visto que el procedimiento más usado en la práctica genera cambios de nivel espurios y que el procedimiento propuesto aventaja al anterior. Finalmente, se ha aplicado el contraste propuesto a la serie de rendimientos diarios del tipo de cambio de la peseta frente al dólar estadounidense, detectándose dos cambios de nivel.

## 5.8 Apéndice A

### Estadístico utilizado por Tsay para detectar cambios de nivel

Considérese la serie de interés  $y_t = \mu + a_t$ , para  $t = 1, \dots, \tau, \dots, T$ , donde  $\mu$  es la media y  $a_t$  es una secuencia de innovaciones i.i.d. que se suponen  $N(0, \sigma_a^2)$ . La serie observada,  $z_t$ , está contaminada con un cambio de nivel de tamaño  $w$  en la observación  $t = \tau$ , es decir,

$$z_t = y_t + wS_t^\tau = \mu + wS_t^\tau + a_t$$

donde  $S_t^\tau = I(t > \tau)$  es la función escalón, que toma valor cero si  $t < \tau$  y uno si  $t \geq \tau$ .

El estadístico propuesto por Chen y Tiao (1986) y generalizado por Tsay (1988) para contrastar la hipótesis nula  $H_0 : w = 0$ , viene dado por:

$$\lambda_\tau = \frac{\hat{w}}{\widehat{Var}(\hat{w})} \quad (5.13)$$

donde  $\hat{w} = \frac{Cov(z_t, S_t^\tau)}{Var(S_t^\tau)}$  y  $\widehat{Var}(\hat{w}) = \frac{\hat{\sigma}_a^2}{TVar(S_t^\tau)}$ . Como  $Cov(z_t, S_t^\tau) = \frac{1}{T} \sum_{t=\tau}^T (z_t - \bar{z})$  y  $Var(S_t^\tau) = \frac{\tau(T-\tau)}{T^2}$ , sustituyendo en (5.13) se tiene que, el estadístico de interés será

$$\lambda_\tau = \frac{\sum_{t=\tau}^T (z_t - \bar{z})}{\hat{\sigma}_a \sqrt{\frac{\tau(T-\tau)}{T}}}$$

donde  $\hat{\sigma}_a^2$  es la varianza de los residuos dada por

$$\hat{\sigma}_a^2 = \frac{1}{T} \sum_{t=1}^T (z_t - \hat{\mu} - \hat{w}S_t^\tau)^2.$$

con  $\hat{\mu} = \bar{z} - \hat{w}\bar{S}$   $\bar{S} = \frac{T-\tau}{T}$ . Después de algunos cálculos, es fácil ver que

$$\hat{\sigma}_a^2 = \frac{1}{T} \sum_{t=1}^T (z_t - \bar{z})^2 - \frac{1}{\tau(T-\tau)} \left[ \sum_{t=\tau}^T (z_t - \bar{z}) \right]^2.$$

5.9 Apéndice de tablas y figuras

Tabla 5.1: Percentiles de la distribución empírica de  $\lambda=\max_{m=1,\dots,T} |\lambda_m|$

Percentil	T=25	T=200	T=500	T=1000	T=5000	T=15000
80%	2.62	2.65	2.73	2.77	2.88	2.91
85%	2.78	2.79	2.84	2.89	3.01	3.03
90%	2.98	2.97	3.01	3.04	3.15	<b>3.18</b>
95%	3.36	3.23	3.26	3.28	3.39	<b>3.43</b>
99%	4.12	3.77	3.78	3.77	3.87	<b>3.90</b>

Figura 5.1: Estimadores kernel de la densidad de  $\lambda$

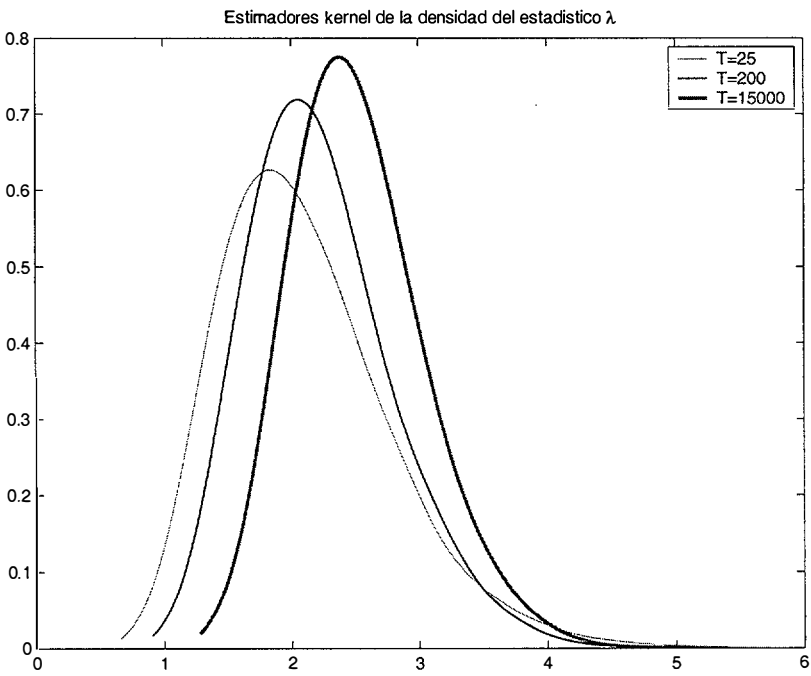


Tabla 5.2: Tamaño del contraste de Tsay bajo heterocedasticidad condicional

		$\alpha$ 0.1 y $C_T$ 3.18			$\alpha$ 0.05 y $C_T$ 3.43			$\alpha$ 0.01 y $C_T$ 3.90		
Modelo	Parámetros	T 500	T 1000	T 5000	T 500	T 1000	T 5000	T 500	T 1000	T 5000
GARCH $(\omega, \alpha, \beta)$	(0.10,0.10,0.80)	0.09	0.10	0.12	0.05	0.05	0.07	0.02	0.02	0.02
	(0.02,0.10,0.88)	0.13	0.14	0.17	0.08	0.09	0.10	0.03	0.04	0.05
	(0.07,0.05,0.88)	0.08	0.08	0.10	0.04	0.04	0.06	0.01	0.01	0.01
	(0.05,0.15,0.80)	0.12	0.13	0.16	0.08	0.09	0.11	0.04	0.04	0.05
GARCH- $t_7$ $(\omega, \alpha, \beta)$	(0.10,0.10,0.80)	0.12	0.13	0.16	0.08	0.09	0.11	0.04	0.04	0.05
	(0.02,0.10,0.88)	0.30	0.34	0.37	0.25	0.29	0.33	0.18	0.23	0.27
	(0.07,0.05,0.88)	0.10	0.11	0.13	0.06	0.07	0.08	0.03	0.03	0.03
	(0.05,0.15,0.80)	0.18	0.19	0.23	0.14	0.15	0.18	0.09	0.10	0.13
EGARCH $(\omega, \alpha, \beta, \gamma)$	(-0.004,0.20,0.95,0.05)	0.10	0.11	0.13	0.06	0.07	0.08	0.02	0.03	0.03
	(-0.001,0.10,0.98,-0.05)	0.10	0.11	0.13	0.06	0.06	0.08	0.02	0.03	0.03
	(-0.004,0.15,0.95,0.10)	0.10	0.12	0.14	0.06	0.07	0.08	0.03	0.03	0.04
	(-0.010,0.30,0.98,-0.10)	0.20	0.21	0.23	0.15	0.16	0.17	0.09	0.09	0.10
ARSV $(\sigma_*^2, \phi, \sigma_\eta^2)$	(0.77,0.90,0.10)	0.16	0.17	0.19	0.10	0.11	0.12	0.04	0.05	0.05
	(0.78,0.95,0.05)	0.18	0.20	0.22	0.12	0.14	0.15	0.05	0.06	0.07
	(0.29,0.98,0.10)	0.26	0.26	0.26	0.20	0.20	0.20	0.12	0.13	0.13
	(0.80,0.98,0.02)	0.20	0.24	0.27	0.14	0.16	0.18	0.06	0.08	0.09





Tabla 5.5: Percentiles de la distribución empírica de  $e=\max_{m=1,\dots,T} |e_m|$

Percentil	T=25	T=200	T=500	T=1000	T=5000	T=15000	Asintótico
80%	0.98	1.03	1.05	1.05	1.07	1.07	1.07
85%	1.04	1.10	1.12	1.12	1.13	1.13	1.14
90%	1.11	1.19	1.20	1.21	1.22	<b>1.22</b>	<b>1.22</b>
95%	1.23	1.32	1.34	1.34	1.35	<b>1.36</b>	<b>1.36</b>
99%	1.44	1.60	1.62	1.62	1.64	<b>1.64</b>	<b>1.63</b>

Figura 5.2: Estimadores kernel de la densidad de  $e$

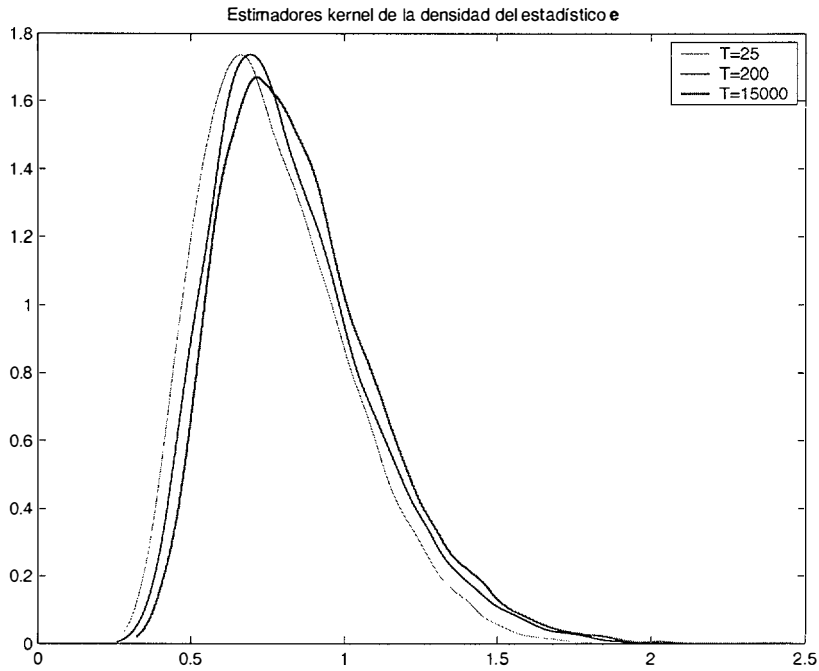


Tabla 5.6: Tamaño del nuevo contraste bajo heterocedasticidad condicional

		$\alpha = 0.1$ y $C_N = 1.22$			$\alpha = 0.05$ y $C_N = 1.36$			$\alpha = 0.01$ y $C_N = 1.63$		
Modelo	Parámetros	T=500	T=1000	T=5000	T=500	T=1000	T=5000	T=500	T=1000	T=5000
GARCH $(\omega, \alpha, \beta)$	(0.10,0.10,0.80)	0.09	0.10	0.10	0.04	0.04	0.05	0.01	0.01	0.01
	(0.02,0.10,0.88)	0.10	0.10	0.10	0.05	0.05	0.05	0.01	0.01	0.01
	(0.07,0.05,0.88)	0.09	0.09	0.10	0.04	0.04	0.05	0.01	0.01	0.01
	(0.05,0.15,0.80)	0.10	0.10	0.10	0.05	0.05	0.05	0.01	0.01	0.01
GARCH- $t_7$ $(\omega, \alpha, \beta)$	(0.10,0.10,0.80)	0.10	0.10	0.10	0.05	0.05	0.05	0.01	0.01	0.01
	(0.02,0.10,0.88)	0.16	0.18	0.22	0.10	0.11	0.15	0.03	0.04	0.06
	(0.07,0.05,0.88)	0.10	0.10	0.10	0.05	0.05	0.05	0.01	0.01	0.01
	(0.05,0.15,0.80)	0.12	0.13	0.14	0.07	0.07	0.08	0.02	0.02	0.02
EGARCH $(\omega, \alpha, \beta, \gamma)$	(0.004,0.20,0.95,0.05)	0.09	0.09	0.10	0.05	0.04	0.05	0.01	0.01	0.01
	(0.001,0.10,0.98,0.05)	0.10	0.09	0.10	0.05	0.04	0.05	0.01	0.01	0.01
	(0.004,0.15,0.95,0.10)	0.10	0.10	0.10	0.05	0.05	0.05	0.01	0.01	0.01
	(0.010,0.30,0.98,0.10)	0.12	0.11	0.10	0.06	0.05	0.06	0.01	0.01	0.01
ARSV $(\sigma_*^2, \phi, \sigma_\eta^2)$	(0.77,0.90,0.10)	0.10	0.10	0.10	0.04	0.04	0.05	0.01	0.01	0.01
	(0.78,0.95,0.05)	0.10	0.10	0.10	0.05	0.05	0.05	0.01	0.01	0.01
	(0.29,0.98,0.10)	0.11	0.10	0.11	0.06	0.05	0.05	0.02	0.01	0.01
	(0.80,0.98,0.02)	0.10	0.10	0.10	0.05	0.05	0.05	0.01	0.01	0.01

Tabla 5.7: Potencia del nuevo contraste cuando  $y_t \sim N(0, 1)$

[illegible]

Tabla 5.8: Potencia del nuevo contraste cuando  $y_t \sim GARCH(1, 1)$

[illegible]

Tabla 5.9: Ratio, en porcentaje, entre los percentiles de la distribución empírica de  $\mathbf{e}=\max_{m=1,\dots,T} |e_m|$  aplicado a los residuos y los percentiles de la tabla 5.5

AR(1) con $\phi = 0.5$						
Percentil	T 25	T=200	T 500	T=1000	T 5000	T 15000
80%	96.94	100.0	100.0	100.9	100.0	100.9
85%	95.19	98.18	99.11	100.0	100.0	100.0
90%	93.69	97.48	100.0	99.17	100.0	100.0
95%	90.24	97.73	99.25	99.25	100.0	100.0
99%	86.11	94.37	96.91	98.15	98.17	99.39
AR(1) con $\phi = 0.9$						
Percentil	T=25	T=200	T=500	T=1000	T=5000	T=15000
80%	103.1	98.06	99.05	100.0	100.0	100.0
85%	101.9	96.36	97.32	98.21	100.0	100.0
90%	100.0	93.28	96.67	97.52	99.18	100.0
95%	97.56	90.15	94.03	97.01	98.52	99.26
99%	94.44	84.38	89.51	93.83	97.56	98.78
AR(3) con $\phi_1 = 0.2, \phi_2 = 0.5$ y $\phi_3 = -0.1$						
Percentil	T=25	T=200	T=500	T=1000	T=5000	T=15000
80%	98.98	98.06	99.05	100.0	100.0	100.0
85%	97.12	96.36	98.21	99.11	100.0	100.0
90%	95.50	95.80	98.33	99.17	100.0	100.0
95%	91.87	94.70	97.76	98.51	100.0	100.0
99%	87.50	90.00	95.06	96.91	98.17	99.39

Tabla 5.10: Comparación de dos procedimientos para detectar dos cambios de nivel con el estadístico  $e$  y  $T = 500$ 

	Procedimiento (D)						Procedimiento (C)					
	% de rechazos de $H_0$			Mediana instantánea			% de rechazos de $H_0$			Mediana instantánea		
$\tau_1 - \tau_2$	Serie original	Subserie 1	Subserie 2	Serie original	Subserie 1	Subserie 2	Serie original	Corregida 1	Corregida 2	Serie original	Corregida 1	Corregida 2
50 – 100	100	98.31	0	100	50	-	100	7.33	6.86	100	50	51
125 – 175	100	98.92	0	174	123	-	100	63.95	57.53	174	123	189
250 – 300	100	24.91	1.25	259	224	329	100	53.10	29.28	259	302	259
375 – 425	100	4.14	57.75	375	180	426	100	48.37	43.49	375	424	348
400 – 450	100	4.17	59.90	399	191	451	100	23.33	20.97	399	449	363
50 – 175	100	99.05	0	174	51	-	100	40.47	36.64	174	52	193
125 – 250	100	99.92	0	247	125	-	100	99.92	98.74	247	125	244
250 – 375	100	8.23	98.90	252	169	374	100	99.98	98.81	252	374	255

Tabla 5.11: Comparación de dos procedimientos para detectar dos cambios de nivel con el estadístico  $e$  y  $T = 1000$ 

	Procedimiento (D)						Procedimiento (C)					
	% de rechazos de $H_0$			Mediana instantánea			% de rechazos de $H_0$			Mediana instantánea		
$\tau_1 - \tau_2$	Serie original	Subserie 1	Subserie 2	Serie original	Subserie 1	Subserie 2	Serie original	Corregida 1	Corregida 2	Serie original	Corregida 1	Corregida 2
100 – 200	100	100	0	200	100	-	100	64.66	60.10	200	100	105
250 – 350	100	100	0	349	249	-	100	99.66	99.36	349	249	365
500 – 600	100	31.18	53.55	513	468	778	100	94.18	84.19	513	602	514
750 – 850	100	4.28	99.49	750	390	851	100	98.74	98.09	750	849	717
800 – 900	100	4.42	98.82	799	398	900	100	87.95	86.10	799	899	750
100 – 350	100	100	0	349	101	-	100	98.58	97.79	349	101	371
250 – 500	100	100	0	497	250	-	100	100	100	497	250	493
500 – 750	100	7.12	100	502	303	748	100	100	100	502	749	506

Figura 5.3: Procedimiento (C): detectar, estimar y corregir. Un ejemplo

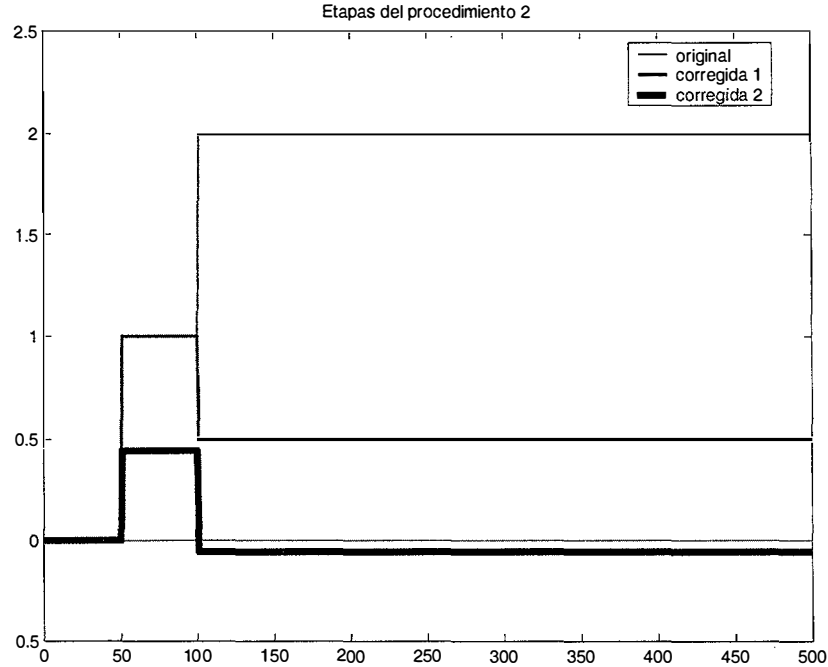


Figura 5.4: Ejemplos

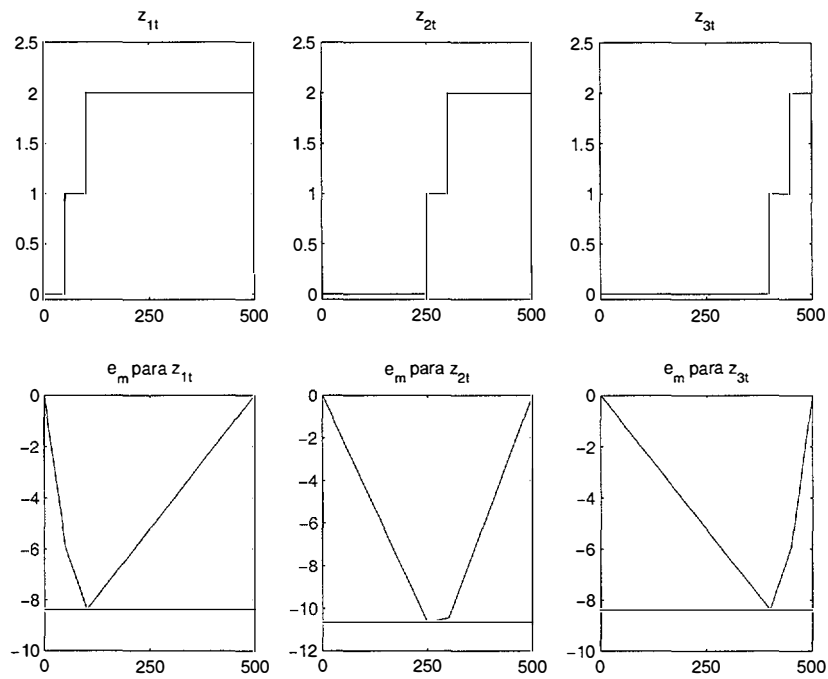


Tabla 5.12: Comparación de dos procedimientos para detectar dos cambios de nivel con el estadístico  $\lambda$  y  $T$  500

$\tau_1 - \tau_2$	Procedimiento (D)						Procedimiento (C)					
	% de rechazos de $H_0$			Mediana instantánea			% de rechazos de $H_0$			Mediana instantánea		
	Serie original	Subserie 1	Subserie 2	Serie original	Subserie 1	Subserie 2	Serie original	Corregida 1	Corregida 2	Serie original	Corregida 1	Corregida 2
50 – 100	100	73.30	24.86	95	50	141	100	47.39	4.17	95	49	97
125 175	100	62.73	43.38	160	125	205	100	28.66	4.04	160	124	168
250 300	100	51.30	56.65	271	250	316	100	21.28	3.16	271	299	280
375 425	100	34.31	68.85	383	374	430	100	40.79	5.97	383	425	378
400 450	100	24.53	74.20	404	399	451	100	48.41	4.91	404	450	397
50 175	100	97.69	6.34	173	50	284	100	98.34	13.79	173	50	173
125 250	100	78.83	34.80	243	125	359	100	99.45	74.39	243	125	237
250 375	100	35.59	78.51	256	250	377	100	99.52	74.58	256	374	263

Tabla 5.13: Comparación de dos procedimientos para detectar dos cambios de nivel con el estadístico  $\lambda$  y  $T$  1000

$\tau_1 - \tau_2$	Procedimiento (D)						Procedimiento (C)					
	% de rechazos de $H_0$			Mediana instantánea			% de rechazos de $H_0$			Mediana instantánea		
	Serie original	Subserie 1	Subserie 2	Serie original	Subserie 1	Subserie 2	Serie original	Corregida 1	Corregida 2	Serie original	Corregida 1	Corregida 2
100 200	100	91.76	22.58	196	100	270	100	96.15	55.72	196	99	194
250 350	100	76.43	49.98	329	250	411	100	87.64	60.82	329	249	332
500 600	100	62.48	67.16	543	500	630	100	82.05	55.15	543	598	539
750 850	100	34.91	85.24	757	749	851	100	93.28	63.43	757	850	758
800 900	100	22.74	91.10	803	799	899	100	96.32	57.66	803	900	805
100 – 350	100	99.82	4.11	348	100	472	100	100	77.86	348	100	348
250 500	100	86.25	28.52	495	250	728	100	100	99.96	495	250	488
500 750	100	29.35	86.14	505	500	750	100	100	99.93	505	749	511

Figura 5.5: Ejemplos

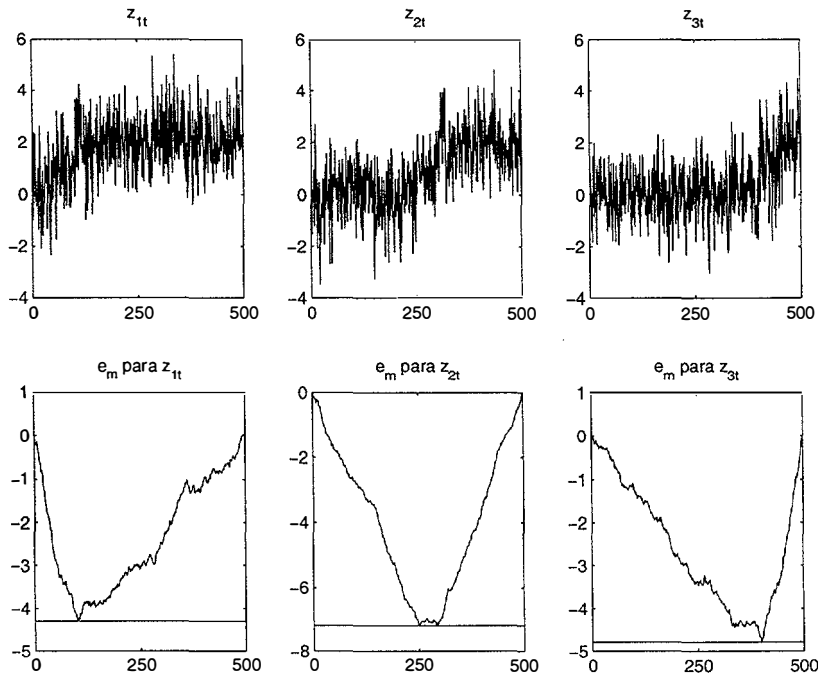


Figura 5.6: Tipo de cambio peseta/dólar

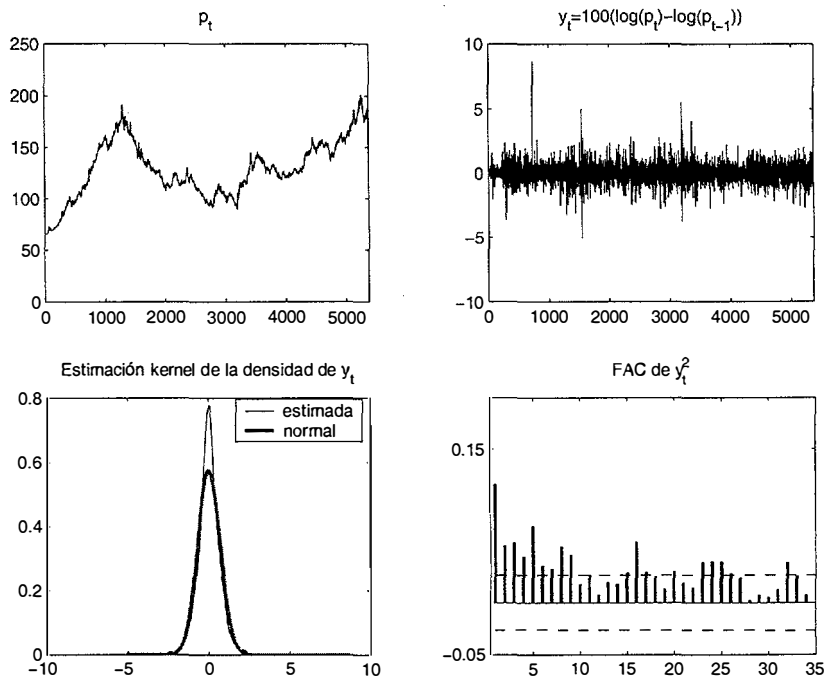




Tabla 5.14: Estadísticos descriptivos de la serie de rendimientos  $y_t$  y de sus subseries

	$y_t$	$y_{1t}$	$y_{2t}$	$y_{3t}$
T	5371	1289	1900	2182
	1 5371	1 1289	1290-3189	3190 5371
Media	0.0195*	0.0826*	0.0388*	0.0330*
D.T.	0.6928	0.6676	0.7285	0.6715
Simetría	0.3684*	1.3940*	0.0636	0.1590*
Curtosis	10.8059*	25.1780*	6.8165*	7.0381*
$r(1)$	0.0260	0.1060*	0.0010	0.0170
$D_{20}(y_t)$	11.77	21.43*	16.03	20.52*
Autocorrelaciones de $y_t^2$				
$r_2(1)$	0.1160*	0.0190	0.2880*	0.1450*
$r_2(2)$	0.0560*	0.0130	0.0880*	0.1100*
$r_2(5)$	0.0750*	0.0170	0.0790*	0.2020*
$r_2(10)$	0.0180	0.0000	0.0370*	0.0440*
$D_{20}(y_t^2)$	129.11*	2.05	179.48*	167.74*

T: Tamaño muestral.

$r(\tau)$ : Autocorrelación de orden  $\tau$  de las observaciones  $y_t$ .

$r_2(\tau)$ : Autocorrelación de orden  $\tau$  de los cuadrados  $y_t^2$ .

$D_{20}(y_t)$  y  $D_{20}(y_t^2)$ : Estadístico propuesto por Peña y Rodriguez (2002) para  $y_t$  e  $y_t^2$  respectivamente (17.41 es el valor crítico al 5%).

\*Significativamente distinto de cero al 95% de confianza.

Figura 5.7: Valores del estadístico  $e$  para la serie  $y_t$  y las subseries

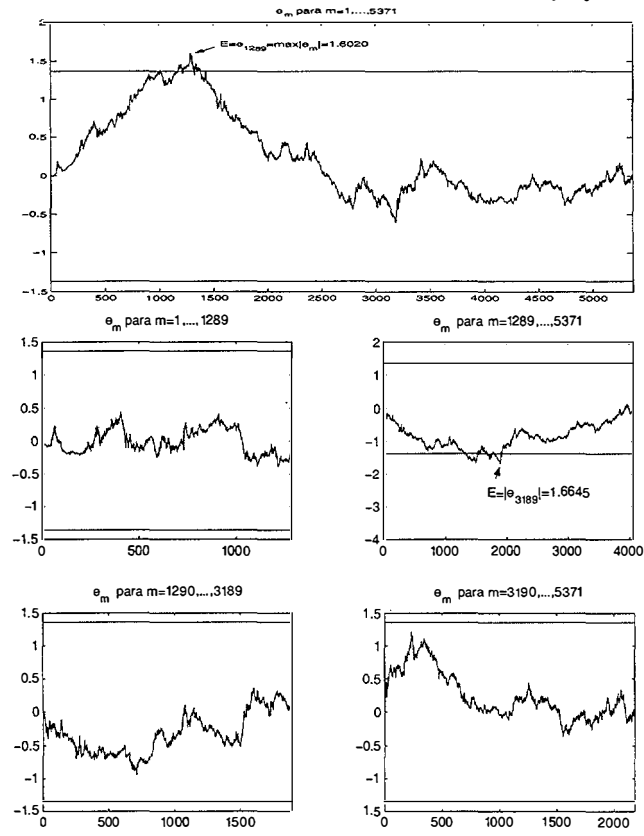
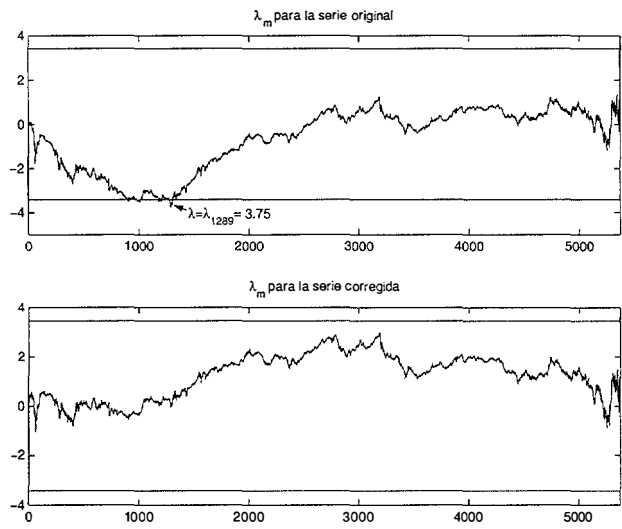


Figura 5.8: Valores del estadístico  $\lambda$  para la serie  $y_t$  y la serie corregida



## Capítulo 6

# Consideraciones finales y líneas de investigación futura

En este trabajo se ha abordado el problema de la modelización de la volatilidad de series temporales financieras. Para ello, se han considerado tres características que aparecen comúnmente en este tipo de datos: (1) heterocedasticidad condicional, (2) observaciones atípicas, y (3) cambios de nivel.

En el Capítulo 2 se ha hecho una revisión de la literatura existente sobre los modelos de heterocedasticidad condicional más utilizados en la práctica para representar series de rendimientos financieros. En particular, se han analizado, más detalladamente, los modelos GARCH y ARSV ilustrando sus características y principales debilidades, mediante el análisis de doce series diarias de rendimientos financieros: siete tipos de cambio y cinco índices bursátiles.

En el Capítulo 3 se ha discutido la relación entre la curtosis, autocorrelación de orden uno de las observaciones al cuadrado y persistencia en modelos GARCH(1,1) y ARSV(1). Se ha visto cómo esta relación puede explicar por qué la persistencia estimada en modelos ARSV(1) es, por lo general, menor que la estimada en modelos

GARCH(1,1) y por qué los modelos ARSV(1) gaussianos parecen adecuados para representar las propiedades de series reales mientras que en los modelos GARCH(1,1), la distribución de las innovaciones, por lo general, debe ser leptocúrtica.

Como investigación futura a partir de los resultados obtenidos en este capítulo, se pretende estudiar como se relacionan los tres coeficientes de interés en los modelos asimétricos. Parece razonable pensar que, si en el modelo EGARCH(1,1), el parámetro de asimetría da más flexibilidad, suceda lo mismo en los modelos de Volatilidad Estocástica asimétricos. Otras futuras extensiones de este capítulo son estudiar el mejor modo de comparar los modelos GARCH y ARSV y caracterizar la persistencia en este tipo de modelos.

En el Capítulo 4 se ha abordado el problema de la presencia simultánea de heterocedasticidad condicional y observaciones atípicas en series temporales. Se han estudiado los efectos de los atípicos en la identificación y estimación de la heterocedasticidad condicional, encontrando: (1) un único atípico sesga hacia cero todos los coeficientes de correlación de las observaciones al cuadrado, y (2) rachas de atípicos generan correlaciones en los cuadrados de la serie. También se han obtenido los sesgos asintóticos causados por los atípicos en los estimadores MCO de procesos ARCH( $p$ ), encontrando que atípicos aislados llevan a infraestimar los parámetros que miden la heterocedasticidad y atípicos consecutivos los sobreestiman. Finalmente, se ha propuesto un procedimiento, en tres pasos, para tratar la presencia simultánea de observaciones atípicas y heterocedasticidad condicional, que consiste en: (i) detectar y corregir los atípicos más extremos con objeto de evitar sesgos elevados en una primera estimación de las varianzas condicionales, (ii) calcular las observaciones estandarizadas, detectar y corregir los atípicos restantes en dichas observaciones siguiendo los métodos tradicionales para series homocedásticas y, (iii) reestimar las varianzas condicionales.

Como futuras extensiones de este capítulo, se pretende estudiar los atípicos en la volatilidad. Todo el análisis en este capítulo se ha hecho para atípicos aditivos que

afectan solamente al nivel de la serie, ya que se sospecha que sus efectos son los peores. Sin embargo, sería también interesante descubrir los efectos que causan los atípicos en la volatilidad. Además, se pretende estudiar también cuál es la mejor manera de corregir las observaciones atípicas una vez que se han detectado. Los métodos tradicionales sustituyen el atípico por la media condicional de la serie, sin embargo al tener varianzas condicionales que cambian en el tiempo no es tan claro que esa sea la mejor manera. Otro problema interesante es estudiar los efectos que causan los atípicos en la estimación de modelos ARCH y GARCH con distribuciones condicionales leptocúrticas. Se pretende estudiar también las implicaciones que la presencia de observaciones atípicas tiene en la predicción de la volatilidad.

En el Capítulo 5 se han estudiado las propiedades del estadístico propuesto por Chen y Tiao (1986) y generalizado por Tsay (1988), comúnmente usado para detectar cambios de nivel en series temporales. Dado que la distribución asintótica de dicho estadístico no es conocida y su distribución en muestras finitas depende crucialmente del tamaño muestral y del proceso generador de los datos, se ha propuesto un nuevo contraste para el que se ha encontrado su distribución asintótica e ilustrado, mediante experimentos de Monte Carlo, sus propiedades en muestras finitas. Se ha propuesto además un procedimiento alternativo para detectar múltiples cambios de nivel en una serie, que muestra claras ventajas con respecto al más utilizado en la práctica. La ventaja fundamental del procedimiento propuesto es que trata de detectar todos los cambios existentes en la serie antes de estimar sus tamaños y así evitar los sesgos en la estimación y la alta probabilidad de generar cambios de nivel espurios.

Una primera extensión de este capítulo es estudiar el comportamiento del contraste propuesto cuando hay varios cambios de nivel de distintos tamaños y signos. La segunda extensión es comparar el contraste propuesto con los propuestos por Balke (1993) y Vogelsang (1998).



# Referencias

- Aggarwal, R., Inclan, C., y Leal, R. (1999). Volatility in Emerging Stock Markets. *Journal of Financial and Quantitative Analysis*, **34**, 33–35.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., y Ebens, H. (2001). The Distribution of Realized Stock Return Volatility. *Journal of Financial Economics*, **61**, 43–76.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., y Labys, P. (2001). The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association*, **96**, 42–55.
- Andersson, J. (2001). On the Normal Inverse Gaussian Stochastic Volatility Model. *Journal of Business and Economic Statistics*, **19**, 44–54.
- Andrews, D.W.K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica*, **61**, 821–856.
- Aparicio, F.M. y Estrada, J. (2001). Empirical Distributions of Stock Returns: European Securities Markets, 1990–95. *The European Journal of Finance*, **7**, 1–21.
- Bai, J. (1997). Estimating Multiple Breaks One at a Time. *Econometric Theory*, **13**, 315–352.
- Bai, X., Russell, J.R., y Tiao, G. (2000). Kurtosis of GARCH and Stochastic Volatility Models. Technical report, Graduate School of Business. University of Chicago. Manuscript.
- Baillie, R.T. y Bollerslev, T. (1989). The Message in Daily Exchange Rates: A Conditional Tale. *Journal of Business and Economic Statistics*, **7**, 297–305.

- Balke, N.S (1993). Detecting Level, Shifts in Time Series. *Journal of Business and Economic Statistics*, **11**, 81–92.
- Balke, N.S. y Fomby, T.B. (1994). Large Shocks, Small Shocks and Economic Fluctuations: Outliers in Macroeconomic Time Series. *Journal of Applied Econometrics*, **9**, 181–200.
- Barndorff-Nielsen, O.E. (1997). Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling. *Scandinavian Journal of Statistics*, **24**, 1–13.
- Barndorff-Nielsen, O.E. y Shephard, N. (2001b). Estimating Quadratic Variation Using Realized Volatility. Technical report, Nuffield College. Oxford.
- Barndorff-Nielsen, O.E. y Shephard, N. (2001a). Non Gaussian Ornstein-Uhlenbeck-Based Models and Some of Their Uses in Financial Economics(with discussion). *Journal of the Royal Statistical Society. Series B*, **63**, 167–241.
- Barndorff-Nielsen, O.E. y Shephard, N. (2002a). Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society. Series B*, **64**. Forthcoming.
- Barndorff-Nielsen, O.E. y Shephard, N. (2002b). *How Accurate is the Asymptotic Approximation to the Distribution of Realized Volatility?* (in D.W.F. Andrews, J.L. Powell, P.A. Rund and J.H. Stock ed.). Identification and Inference for Econometric Models.
- Bauwens, L., Lubrano, M., y Richard, J.F. (1999). *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press.
- Bera, A.K. y Higgins, M.L. (1995). On ARCH Models: Properties, Estimation and Testing. In: *Surveys in Econometrics*, L. Oxley, D.A.R. George, C.J. Roberts and S. Sayer (eds). Oxford:Blackwell.
- Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons.
- Black, F. (1976). Studies of Stock Price Volatility Changes. *Proceedings of the 1976 Meetings of the Business and Economic Statistics, American Statistical Society*,



177–181.

- Blanco, R. (2000). Efectos sobre la Volatilidad del Mercado Bursátil de la Introducción de los Contratos de Futuros y Opciones sobre el IBEX35. *Investigaciones Económicas*, **24**, 139–175.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.
- Bollerslev, T. (1987). A Conditionally Heteroskedasticity Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics*, **69**, 542–547.
- Bollerslev, T. (1988). On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Process. *Journal of Time Series Analysis*, **9**, 121–131.
- Bollerslev, T., Chou, R.Y., y Kroner, K.F. (1992). ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, **52**, 498–505.
- Bollerslev, T. y Engle, R.F. (1993). Common Persistence in Conditional Variances. *Econometrica*, **61**, 166–187.
- Bollerslev, T., Engle, R. F., y Nelson, D. B. (1994). ARCH models. *The Handbook of Econometrics*, **4**, 2959–3038.
- Bollerslev, T. y Mikkelsen, H.O. (1996). Modeling and Pricing Long-Memory in Stock Market Volatility. *Journal of Econometrics*, **73**, 151–184.
- Bougerol, P. y Picard, N. (1992). Stationarity of GARCH Processes and Some Non-negative Time Series. *Journal of Econometrics*, **52**, 115–128.
- Breidt, F.J., Davis, R.A., y Trincade, A. (2001). Least Absolute Deviation Estimation for All-Pass Time Series Models. *Annals of Statistics*, **29**, 919–946.
- Brooks, R.D., Faff, R.W., McKenzie, M.D., y Mitchell, H. (2000). A Multi-Country Study of Power ARCH Models and National Stock Market Returns. *Journal of International Money and Finance*, **19**, 377–397.

- Campbell, J.Y. y Hentchel, L. (1992). No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns. *Journal of Financial Economics*, **31**, 281–318.
- Campbell, J.Y., Lo, A.W., y MacKinlay, A.C. (1997). *The Econometrics of Financial Markets*. Princeton: Princeton University Press.
- Cao, C.Q. y Tsay, R.S. (1992). Nonlinear Time Series Analysis of Stock Volatilities. *Journal of Applied Econometrics*, **7**, 165–185.
- Chan, W. (1995). Understanding the Effect of Time Series Outliers on Sample Autocorrelations. *Test*, **4**, 179–186.
- Chang, I., Tiao, G.C., y Chen, C. (1988). Estimation of Time Series Parameters in the Presence of Outliers. *Technometrics*, **30**, 193–204.
- Chen, C. y Liu, J. (1993). Joint Estimation of Model Parameters and Outlier Effects in Time Series. *Journal of the American Statistical Association*, **88**, 284–296.
- Chen, C. y Tiao, G.C. (1986). Some Diagnostic Statistics in Time Series. Technical report, University of Chicago, Statistical Research Center.
- Chen, C. y Tiao, G.C. (1990). Random Level-Shift Time Series Models, ARIMA Approximations, and Level Shift Detection. *Journal of Business and Economic Statistics*, **8**, 83–97.
- Chib, S., Nardari, F., y Shephard, N. (1998). Markov Chain Monte Carlo Methods for Generalized Stochastic Volatility Models. W.p., John M. Olin School of Business, Washington University.
- Chong, C.W., Ahmad, M.I., y Abdullah, M.Y. (1999). Performance of GARCH Models in Forecasting Stock Market Volatility. *Journal of Forecasting*, **18**, 333–343.
- Csorgo, M. y Horvath (1997). *Limit Theorems in Change-Point Analysis*. John Wiley & Sons.
- Deutsch, S.J., Richards, J.E., y Savin, J.J. (1990). Effects of a Single Outlier on ARMA Identification. *Communications in Statistics, Theory and Methods*, **19**, 2207–2227.

- Diebold, F.X. y López, J.A. (1995). ARCH Models. In: *Macroeconomics: Developments, Tensions and Prospects*, K. Hoover (eds). Oxford:Blackwell.
- Ding, Z., Granger, C.W.J., y Engle, R.F. (1993). A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance*, **1**, 83–106.
- Doornik, J.A. (2001). *Ox: An Object-Oriented Matrix Language* (4th ed.). London: Timberlake Consultants Press.
- Duan, J.C. (1997). Augmented GARCH(p,q) Process and its Diffusion Limit. *Journal of Econometrics*, **79**, 97–127.
- Engle, R.F., Hendry, D.F., y Trumble, D. (1985). Small Sample Properties of ARCH Estimators and Tests. *Canadian Journal of Economics*, **18**, 66–93.
- Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, **50**, 987–1007.
- Engle, R. F. y Bollerslev, T. (1986). Modelling the Persistence of Conditional Variances. *Econometric Reviews*, **5**, 1–50.
- Engle, R. F. y Mustafa, C. (1992). Implied ARCH Models for Option Prices. *Journal of Econometrics*, **52**, 289–311.
- Engle, R. F. y Ng, V. K. (1993). Measuring and Testing the Impact of News on Volatility. *Journal of Finance*, **48**, 1749–1778.
- Fama, E. (1965). The Behaviour of Stock Prices. *Journal of Business*, **38**, 34–105.
- Fornani, F. y Mele, A. (1997). Weak Convergence and Distributional Assumptions for a General Class of Nonlinear ARCH models. *Econometric Reviews*, **16**, 205–227.
- Forsberg, L. y Bollerslev, T. (2002). Bridging the Gap Between the Distribution of Realized (ECU) Volatility and ARCH Modeling (of the Euro): A Note on the GARCH-NIG Model. Manuscript.
- Fox, A.J. (1972). Outliers in Time Series. *Journal of the Royal Statistical Society B*, **34**, 350–363.

- Franses, P.H. y Ghijssels, H. (1999). Additive Outliers, GARCH and Forecasting Volatility. *International Journal of Forecasting*, **15**, 1–9.
- Geweke, J. (1986). Comment: Modelling the Persistence of Conditional Variances. *Econometric Reviews*, **5**, 57–61.
- Ghysels, E., Harvey, A., y Renault, E. (1996). Stochastic Volatility. *Handbook of Statistics*, **14**.
- Glosten, L.R., Jagannathan, R., y Runkle, D.E. (1993). On the Relation between Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, **48**, 1779–1801.
- González-Rivera, G. (1998). Smooth Transition GARCH Models. *Studies in Nonlinear Dynamics and Econometrics*, **3**, 61–78.
- Granger, C.W.J. y Ding, Z. (1995). Some Properties of Absolute Return, an Alternative Measure of Risk. *Annales d'Économie et de Statistique*, **40**, 67–90.
- Gregory, A.W. y Reeves, J.J. (2001). Estimation and Inference in ARCH Models in the Presence of Outliers. Technical report, Queen's University.
- Hafner, C.M. y Herwartz, H. (2000). Testing for Linear Autoregressive Dynamics under Heteroskedasticity. *Econometrics Journal*, **3**, 177–197.
- Harvey, A., Ruiz, E., y Shephard, N. (1994). Multivariate Stochastic Variance Models. *Review of Economic Studies*, **61**, 247–264.
- He, C. y Teräsvirta, T. (1997). Statistical Properties of the Asymmetric Power ARCH Process. W.P. 199, Stockholm School of Economics.
- He, C. y Teräsvirta, T. (1999). Properties of Moments of a Family of GARCH Processes. *Journal of Econometrics*, **92**, 173–192.
- He, C., Teräsvirta, T., y Malmsten, H. (1999). Fourth Moment Structure of a Family of First-Order Exponential GARCH Models. W.P. 345, Stockholm School of Economics.

- Hentschel, L. (1995). All in the Family: Nesting Symmetric and Asymmetric GARCH Models. *Journal of Financial Economics*, **39**, 71–104.
- Higgins, M.L. y Bera, A.K. (1992). A Class of Nonlinear ARCH Models. *International Economic Review*, **33**, 137–158.
- Hinkley, D.V. (1971). Inference About the Change-Point from Cumulative Sum Tests. *Biometrika*, **58**, 509–523.
- Hotta, L. K. y Tsay, R. S. (1998). Outliers in GARCH Processes. W.p., The Graduate School of Business, University of Chicago.
- Hsieh, D.A. (1989). Modelling Heteroskedasticity in Daily Foreign Exchange Rates. *Journal of Business and Economic Statistics*, **7**, 307–317.
- Hwang, S. y Satchell, S.E. (2000). Market Risk and the Concept of Fundamental Volatility: Measuring Volatility across Asset and Derivative Markets and Testing for the Impact of Derivatives Markets on Financial Markets. *Journal of Banking and Finance*, **24**, 759–785.
- Inclan, C. y Tiao, G.C. (1994). Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance. *Journal of the American Statistical Association*, **89**, 913–923.
- Jiang, J., Zhao, Q., y Hui, Y.V. (2001). Robust Modelling of ARCH Models. *Journal of Forecasting*, **20**, 111–133.
- Jorion, P. (1988). On Jump Processes in the Foreign Exchange and Stock Markets. *Review of Financial Studies*, **1**, 427–445.
- Kearns, P. y Pagan, A.R. (1993). Australian Stock Market Volatility. *The Economic Record*, **69**, 163–178.
- Kennedy, D.P. (1976). The Distribution of the Maximum Brownian Excursion. *Journal of Applied Probability*, **13**, 371–376.
- Kim, S., Shephard, N., y Chib, S. (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Model. *Review of Economic Studies*, **65**, 361–393.

- Kleibergen, F. y van Dijk, H.K. (1993). Non-stationarity in GARCH Models: A Bayesian Analysis. *Journal of Applied Econometrics*, **8**, 41–61.
- Kokoszka, P. y Leipus, R. (1998). Change-Point in the Mean of Dependent Observations. *Statistics & Probability Letters*, **40**, 385–393.
- Koopman, S.J., N. Shephard y Doornik, J. (1999). Statistical Algorithms for Models in State Space using Ssfpack 2.2. *Econometrics Journal*, **2**, 113–166.
- Lee, J.H.H. (1991). A Lagrange Multiplier Test for GARCH Models. *Economics Letters*, **37**, 265–271.
- León, A. y Mora, J. (1999). Modelling Conditional Heteroskedasticity: Application to the IBEX-35 Stock-Return Index. *Spanish Economic Review*, **1**, 215–238.
- Li, W.K., Ling, S., y McAleer, M. (2001). A Survey of Recent Theoretical Results for Time Series Models with GARCH Errors. Manuscript.
- Liesenfeld, R. y Jung, R.C. (2000). Stochastic Volatility Models: Conditional Normality versus Heavy-Tailed Distributions. *Journal of Applied Econometrics*, **15**, 137–160.
- Ling, S. y McAleer, M. (2002a). Necessary and Sufficient conditions for the GARCH(r,s) and Asymmetric GARCH(r,s) Models. *Econometric Theory*, **18**, 722–729.
- Ling, S. y McAleer, M. (2002b). Stationarity and the Existence of moments of a family of GARCH processes. *Journal of Econometrics*, **106**, 109–117.
- Lobato, I. y Savin, N.E. (1998). Real and Spurious Long-Memory Properties of Stock-Market Data. *Journal of Business and Economic Statistics*, **16**, 475–495.
- Lobato, I. y Velasco, C. (2000). Long Memory in Stock-Market Trading Volume. *Journal of Business and Economic Statistics*, **18**, 410–427.
- Loudon, G.F., Watt, W.H., y Yadav, P.K. (2000). An Empirical Analysis of Alternative Parametric ARCH Models. *Journal of Applied Econometrics*, **15**, 117–136.
- Lumsdaine, R.L. y Ng, S. (1999). Testing for ARCH in the Presence of a Possibly Misspecified Conditional Mean. *Journal of Econometrics*, **93**, 257–279.

- Mahieu, R.J. y Schotman, P.C. (1998). An Empirical Application of Stochastic Volatility Models. *Journal of Applied Econometrics*, **13**, 333–360.
- Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, **36**, 394–419.
- McCurdy, T.H. y Morgan, I. (1987). Tests of the Martingale Hypothesis for Foreign Currency Futures with Time Varying Volatility. *International Journal of Forecasting*, **3**, 131–148.
- McLeod, A.J. y Li, W.K. (1983). Diagnostic Checking ARMA Time Series Models Using Squared-Residual Correlations. *Journal of Time Series Analysis*, **4**, 269–273.
- Mendes, B.V.M. (2000). Assessing the Bias of Maximum Likelihood Estimates of Contaminated GARCH Models. *Journal of Statistical Computation and Simulation*, **67**, 359–376.
- Milhoj, A. (1987). A Conditional Variance Model for Daily Observations of an Exchange Rate. *Journal of Business and Economic Statistics*, **5**, 99–103.
- Mills, T.C. (1996). Non-Linear Forecasting of Financial Time Series: An Overview and some New Models. *Journal of Forecasting*, **15**, 127–135.
- Muler, N. y Yohai, V. (2002). Robust Estimates for ARCH Processes. *Journal of Time Series Analysis*, **23**, 341–375.
- Nelson, D.B. (1988). *Time Series Behavior of Stock Market Volatility and Returns*. Ph. D. thesis, M.I.T.
- Nelson, D. B. (1990). Stationary and Persistence in the GARCH(1,1) Model. *Econometric Theory*, **6**, 318–344.
- Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, **59**, 347–370.
- Nelson, D. B. y Cao, C.Q. (1992). Inequality Constraints in the Univariate GARCH Model. *Journal of Business and Economic Statistics*, **10**, 229–235.

- Nelson, D. B. y Foster, D.P. (1994). Asymptotic Filtering Theory for Univariate ARCH Models. *Econometrica*, **62**, 1–41.
- Palm, F.C. (1997). GARCH Models of Volatility. *Handbook of Statistics*, **14**.
- Pantula, S.G. (1986). Comment: Modelling the Persistence of Conditional Variances. *Econometric Reviews*, **5**, 71–73.
- Paoletta, M.S. (2000). Generalized GARCH Models with Nonnormal Distributions. Technical report, Institute of Statistics and Econometrics, Kiel University. Manuscript.
- Peña, D. (2001). *Outliers, Influential Observations and Missing Data* (en A course in time series. Peña, D., G.C. Tiao and R.S. Tsay ed.). John Wiley, New York.
- Peña, D. y Rodríguez, J. (2002). A Powerful Portmanteau Test of Lack of Fit in Time Series. *Journal of the American Statistical Association*, **97**. Forthcoming.
- Pérez, A. y Ruiz, E. (2002). Properties of the sample autocorrelations of non-linear transformations in Long Memory Stochastic Volatility models. Manuscript.
- Poon, S.H. y Taylor, S.J. (1992). Stock Returns and Volatility: An Empirical Study of the UK Stock Market. *Journal of Banking and Finance*, **16**, 37–59.
- Poterba, J. y Summers, L. (1986). The Persistence of Volatility and Stock Market Fluctuations. *American Economic Review*, **76**, 1124–1141.
- Sakata, S. y White, H. (1998). High Breakdown Point Conditional Dispersion Estimation with Application to S&P 500 Daily Returns Volatility. *Econometrica*, **66**, 529–567.
- Sandmann, G. y Koopman, S.J. (1998). Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood. *Journal of Econometrics*, **87**, 271–301.
- Schwert, G.W. (1989). Why does Stock Market Volatility change over Time? *Journal of Finance*, **45**, 1129–1155.
- Sentana, E. (1995). Quadratic ARCH Models. *Review of Economic Studies*, **62**, 639–661.



- Shephard, N. (1996). Statistical Aspects of ARCH and Stochastic Volatility. In: *Time Series Models in Econometrics, Finance and Other Fields*, D.R. Cox, D.V. Hinkley and O.E. Bardorff-Nielsen (eds). Chapman & Hall, London.
- So, Mike K.P., Lam, K., y Li, W.K. (1999). Forecasting Exchange Rate Volatility using Autoregressive Random Variance Model. *Applied Financial Economics*, **9**, 583–591.
- Taylor, S.J. (1986). *Modelling Financial Time Series*. John Wiley.
- Taylor, S.J. (1994). Modelling Stochastic Volatility: A Review and Comparative Study. *Mathematical Finance*, **4**, 183–204.
- Törnqvist, L., Vartia, P., y Y.O., Vartia (1985). How Should Relative Changes Be Measured? *The American Statistician*, **39**, 43–46.
- Teräsvirta, T. (1996). Two Stylized Facts and the GARCH(1,1) Model. W.P. Series in Finance and Economics 96, Stockholm School of Economics.
- Tolvi, J. (2000). Outliers in Time Series: A Review. Manuscript.
- Tsay, R.S. (2002). *Analysis of Financial Time Series*. Wiley Series in Probability and Statistics.
- Tsay, R. S. (1987). Conditional Heteroscedastic Time Series Models. *Journal of the American Statistical Association*, **82**, 590–604.
- Tsay, R. S. (1988). Outliers, Level Shifts and Variance Changes in Time Series. *Journal of Forecasting*, **7**, 1–20.
- Vaage, K. (2000). Detection of Outliers and Level Shifts in Time Series: An Evaluation of Two Alternative Procedures. *Journal of Forecasting*, **19**, 23–37.
- Van Dijk, D., Frances, P. H., y Lucas, A. (1999). Testing for ARCH in the Presence of Additive Outliers. *Journal of Applied Econometrics*, **14**, 539–562.
- Verhoeven, P. y McAleer, M. (2000). Modelling Outliers and Extreme Observations for ARMA-GARCH Processes. Manuscript.

- Vogelsang, T.J. (1998). Testing for a Shift in Mean without having to estimate Serial-Correlation Parameters. *Journal of Business and Economic Statistics*, **16**, 73–80.
- Vogelsang, T.J. (1999). Sources of Nonmonotonic Power when Testing for a Shift in Mean of a Dynamic Time Series. *Journal of Econometrics*, **88**, 283–299.
- Worsley, K.J. (1986). Confidence Regions and Tests for a Change-Point in a Sequence of Exponential Family Random Variables. *Biometrika*, **73**, 91–104.
- Yang, M. y Bewley, R. (1995). Moving Average Conditional Heteroskedastic Processes. *Economics Letters*, **49**, 367–372.
- Zakoian, J.M. (1994). Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control*, **18**, 931–955.

# **HETEROCEDASTICIDAD CONDICIONAL, ATÍPICOS Y CAMBIOS DE NIVEL EN SERIES TEMPORALES FINANCIERAS**

Doctoranda: María Angeles Carnero Fernández

Directores: Daniel Peña y Esther Ruiz

Octubre 2002

Resumen

En esta Tesis se analizan tres características empíricas frecuentes en las series financieras: heterocedasticidad condicional, observaciones atípicas y cambios de nivel. En primer lugar, se analiza la relación entre el coeficiente de curtosis, la persistencia de las innovaciones sobre la volatilidad y la correlación de orden uno de las observaciones al cuadrado en los modelos GARCH(1,1) y ARSV(1). Dicha relación permite explicar por qué, para las mismas series, la persistencia se estima, a menudo, mayor en modelos GARCH que en ARSV. También se explica por qué los modelos ARSV gaussianos parecen adecuados para explicar las propiedades de las series reales mientras que si se ajusta un modelo GARCH, la distribución condicional necesita ser leptocúrtica. A continuación, se estudia la presencia simultánea de observaciones atípicas y heterocedasticidad condicional en series temporales. Se investigan los efectos causados por atípicos en el diagnóstico y estimación de la heterocedasticidad condicional y se propone un procedimiento para tratar este problema. Por último, se propone un nuevo contraste para detectar cambios de nivel en series temporales, que es robusto a distribuciones marginales no Gaussianas y, en particular, a la presencia de heterocedasticidad condicional. Se propone también un procedimiento para el caso de múltiples cambios de nivel. Todos los resultados se ilustran con diferentes series financieras.

